AFOSR - TR - 72 - 1054

1

MINIMAX PROBLEMS, SADDLE-FUNCTIONS AND DUALITY

by

LYNN McLINDEN

A dissertation submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF WASHINGTON

1971

Approved by ______________________
(Chairman of Supervisory Committee)

Department ______ Mathematics ______
(Departmental Faculty sponsoring candidate)

Date ______ May 28, 1971 ______

AFOSR 71-1994

117

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of Washington Department of Mathematics Seattle, Washington 98195	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

MINIMAX PROBLEMS, SADDLE-FUNCTIONS AND DUALITY

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

Lynn McLinden

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
May 28, 1971	112	60

8a. CONTRACT OR GRANT NO [redacted]-AFOSR-71-1994 b. PROJECT NO 9749 c. 61102F d. 681304	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR - TR - 72 - 1054

10. DISTRIBUTION STATEMENT

A. Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES TECH, OTHER	12. SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research(NM) 1400 Wilson Blvd. Arlington, Virginia 22209

13. ABSTRACT

Minimax problems are fundamented to nonlinear programming, because of the way constraints can be represented using Lagrange multipliers. Better ways of solving minimax problems would lead thus lead to breakthroughs in sovling most other problems of optimization. This dissertation opens a new avenue to the study of minimax problems by developing a theory of dual operations on saddle-functions convex-concave functions parallel to that already known for (purely) convex functions. Results are thereby obtained concerning minimax problems which are dual to each other. It is expected that these results will find computational applications analogous to those already acclaimed in the convex case, for instance in decomposition of large-scale problems.

DD FORM 1 NOV 65 1473

Security Classification

CONTENTS

Acknowledgments

To Professor R. Tyrrell Rockafellar, my thesis advisor, I wish to express my gratitude for introducing me to convex analysis and for suggesting the problem which led to this thesis. Deep appreciation also goes to [redacted] my wife, for her steadfast encouragement. I owe special thanks to the Air Force Office of Scientific Research for its financial support under grant no. AFOSR-71-1994 and to the National Defense Education Act for a generous fellowship at the University of Washington.

Preface

Due to the frequency with which results from Rockafellar [44] are cited throughout the thesis, a special abbreviation is used. Namely, the number of the result being cited is given enclosed in parentheses. For example, Theorem 23.8 is cited as (23.8), Corollary 6.3.1 as (6.3.1), and so forth.

Throughout the thesis expressions sometimes appear which involve taking the supremum or infimum of an empty set of numbers. Whenever these occur they are to be interpreted using the conventions $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.

The common abbreviation "iff" is used for the phrase "if, and only if."

Finally, for §1 the reader need only read §0 up to Lemma 0.3. The rest of the thesis draws on all of §0.

Introduction

Minimax theory may be said to have originated in 1928 with von Neumann's minimax theorem for matrix games [34]. Various proofs and generalizations of this theorem have been given by many authors, including Ville [59], Kakutani [26], Wald [60], Shiffman [50], Fan [18, 19], Kneser [27], Glicksberg [24], Nikaido [35], Berge [4], Sion [51], Ghouila-Houri [23], Moreau [33], and Rockafellar [39, 40].

Much of the early work in minimax theory was done in connection with game theory. However in about 1950 two equivalences were established which made it apparent that minimax theory had much relevance for mathematical programming. One of these equivalences was that between matrix games and dual pairs of linear programs (see Dantzig [12], Gale-Kuhn-Tucker [22], and Charnes [8]). The other equivalence was that between convex programs and Lagrangian saddle-point problems (see Kuhn-Tucker [28], Slater [52], and extensions given by Hurwicz-Uzawa in [2]). Various authors, including Stoer [53, 54], Mangasarian-Ponstein [31], and Dantzig-Eisenberg-Cottle [13], later derived duality results for constrained maximization and minimization problems by means of minimax theorems.

In [39] Rockafellar defined a conjugacy correspondence among saddle-functions parallel to that of Fenchel [20] for convex functions. This correspondence was used in [43] to represent (in finitely many different ways) a certain dual pair of convex programs as a dual pair of minimax problems. At a later date Tynjanskii [57] independently defined the conjugacy correspondence for a more restrictive class of saddle-functions. He used it to associate with a given concave-convex game another game of the same type, and showed how solving such a pair of "dual games" is equivalent to solving a related pair of convex programs. Also, papers of Moreau [33] and Ioffe-Tikhomirov [25] contain implicit results concerning the conjugacy

correspondence among saddle-functions.

The relevance of minimax theory to mathematical economics has long been recognized, dating back to the beginnings of game theory. More recently, minimax theory has been useful in the calculus of variations and optimal control theory (e.g. Rockafellar [47, 48]). It also plays a role in differential games (e.g. Sakawa [49]; see also the survey article by Berkovitz [5]).

Related to minimax problems are max-min problems, i.e. two-stage problems of the form $\max_x(\min_y f(x,y))$. These have been studied by Pshenichnyi [36], Danskin [10], and Bram [6]. Such problems correspond to "half" a saddle-point problem and arise from such practical considerations as two-stage resource allocation.

The preceding references deal primarily with theory. However the task of actually finding saddle-points has also been studied. Work in the early 1950's was done by Brown—von Neumann [7], Robinson [37], and Danskin [9]. Charnes [8] showed that a minimax problem corresponding to a constrained matrix game is equivalent to a dual pair of linear programs, so that such techniques as the simplex method could be applied. Conversely, in order to utilize the Kuhn—Tucker theorem [28] and its generalizations for solving concave programs, Arrow-Hurwicz [2, p. 118] developed a "steepest descent" method for locating the saddle-points of the Lagrangian. More recently, algorithms have been given by Dem'janov [14, 16], Auslender [3], and Danskin [11]. See also Tremolieres' survey paper [56]. Algorithms dealing with max-min problems have been given by Pshenichnyi [36], Dem'janov [15], and Danskin [11].

The problem of minimizing a convex function subject to constraints has been analyzed by various authors by means of the duality theory arising from Fenchel's conjugacy correspondence. This approach, as expounded in [44],

rests ultimately on the duality between two operations which combine a convex function with a linear transformation. The aim of this thesis is to analyze constrained minimax problems in a similar fashion by means of the duality theory arising from the conjugacy correspondence among saddle-functions. To accomplish this we develop for saddle-functions analogues of these fundamental operations on convex functions. But before actually describing our results, we shall sketch the two operations and the applications of them which this thesis extends.

The simpler of the two operations is to form the composition fA of a convex function f with a linear transformation A. The other operation may be called "taking the image of f under A." The resulting function Af is defined by $(Af)(x) = \inf\{f(y) \mid Ay = x\}$. The fundamental result connecting these operations is that, under certain mild hypotheses,

$$(Af)^* = f^*A^*, \tag{1}$$

where $*$ of a linear transformation denotes the adjoint linear transformation and $*$ of a convex function denotes the conjugate convex function.

One of the main consequences of the duality formula (1) is the duality between the operations of addition and infimal convolution for convex functions. This can be obtained by taking f to be the separable function

$$f(x_1,\ldots x_m) = f_1(x_1) + \ldots + f_m(x_m),$$

where each f_i is convex on R^n, and defining A to be the linear transformation which sends each element x of R^n into the m-tuple $(x,\ldots,x)$. In this event fA is $f_1 + \ldots + f_m$ and A^*f^* is the function

$$x^* \to \inf\{f_1^*(x_1^*) + \ldots + f_m^*(x_m^*) \mid x^* = x_1^* + \ldots + x_m^*\},$$

i.e. the infimal convolution of $f_1^*,\ldots,f_m^*$. Formula (1) then implies that under mild hypotheses "the conjugate of the sum is the infimal convolute of the conjugates." This gives a framework encompassing problems of the form

"minimize $h(x)$ subject to $x \in C$, where h and C are convex." Simply take $m = 2$, let $f_1 = h$, and let $f_2(x)$ equal 0 when $x \in C$ and $+\infty$ otherwise.

The duality represented by formula (1) is also fundamental in the perturbational duality theory developed by Rockafellar for generalized convex programs [44]. Among other things, this theory generalizes the classical results about dual linear programs and generalizes Fenchel's Duality Theorem [21, p. 108] (see also [41, 42] and Stoer-Witzgall [55]). It also sheds light on the Lagrange multiplier principle for convex programming and thereby on the celebrated Dantzig-Wolfe decomposition principle for linear and convex programs [44, pp. 285-290] (see also Falk [17] and Lasdon [29]).

In the next three paragraphs we indicate an essential difference between minimax theory and convex function theory, and briefly review the notions of "closed" and "conjugate" for saddle-functions.

The principal difference between convex function theory and minimax theory is not the difference between one and two arguments. Rather it is that in convex function theory the natural object of study is the individual convex function, whereas in minimax theory the natural object of study is a whole equivalence class of saddle-functions. This stems from the fact that there is an equivalence relation among saddle-functions with the property that equivalent saddle-functions have the same (lower and upper) saddle-values and also the same saddle-points. The relation, introduced in [39], is the following: two concave-convex functions K and L are said to be equivalent if and only if the closures of the convex functions $K(x,\cdot)$ and $L(x,\cdot)$ coincide for each x and the closures of the concave functions $K(\cdot,y)$ and $L(\cdot,y)$ coincide for each y.

Recall that in convex function theory, in order to have the crucial

formula

$$(f^*)^* = f \tag{2}$$

hold, one considers convex functions which are lower-semi-continuous, i.e. closed. Similarly, in saddle-function theory one considers "regularized" saddle-functions in order for an analogue of formula (2) to hold. A saddle-function K defined to be closed if and only if it is equivalent to both its concave closure and its convex closure, where by concave (resp. convex) closure we mean the saddle-function obtained from K by closing it (in the sense of convex function theory) in its concave (resp. convex) argument. It is easily seen that a saddle-function is closed if and only if every member of its equivalence class is closed. It is shown in [39] that the property of being a closed saddle-function is constructive. In [39] it is also shown that equivalent closed saddle-functions must be very nearly equal. Roughly speaking, they can differ essentially only at the "corner points" of their "domain of finiteness." In [39] it is shown that each equivalence class $[K]$ of closed saddle-functions is an "interval" in the sense that there exist unique members $\underline{K}$ and $\overline{K}$ of $[K]$ such that $[K]$ contains all, and only those, saddle-functions $\tilde{K}$ satisfying $\underline{K} \leq \tilde{K} \leq \overline{K}$.

We now review the conjugacy correspondence for saddle-functions, first developed in [39]. If K is a concave-convex function from $R^m \times R^n$ to $[-\infty,+\infty]$, the lower conjugate $\underline{K}^*$ and upper conjugate $\overline{K}^*$ of K are defined by

$$\underline{K}^*(x^*,y^*) = \sup_y \inf_x \{\langle x,x^*\rangle + \langle y,y^*\rangle - K(x,y)\}$$

and

$$\overline{K}^*(x^*,y^*) = \inf_x \sup_y \{\langle x,x^*\rangle + \langle y,y^*\rangle - K(x,y)\}.$$

These functions are concave-convex. If K is closed, then $\underline{K}^*$ and $\overline{K}^*$ are equivalent and closed, and moreover they depend only on the equivalence class

[K] containing K. Thus, associated with [K] is a well-defined equivalence class $[K^*]$ of closed concave-convex functions, namely the class containing $\underline{K}^*$ and $\overline{K}^*$. The class $[K^*]$ is said to be the conjugate of [K]. This conjugacy correspondence has the property that the conjugate of $[K^*]$ is [K]. This is the analogue of formula (2) for saddle-functions.

With this review of general facts in mind, we now describe the results obtained in this thesis. We begin with the analogues of the two fundamental operations described above. Let K be a closed concave-convex function and let A be the linear transformation $A_1 \times A_2$ obtained from two other linear transformations A_1 and A_2 by $A_1 \times A_2\,(x,y) = (A_1x, A_2y)$. One of our operations consists of forming an equivalence class [KA] containing all the saddle-functions of the form

$$(x,y) \to \tilde{K}A(x,y) = \tilde{K}(A_1x, A_2y)$$

for $\tilde{K}$ any member of [K]. A mild hypothesis is given which ensures that in fact such a single class exists and, moreover, that all its members are closed. The other operation is to form a single equivalence class [AK] containing all the saddle-functions both of the form

$$(u,v) \to \sup_{\{x|A_1x = u\}} \ \inf_{\{y|A_2y = v\}} \tilde{K}(x,y)$$

and of the form

$$(u,v) \to \inf_{\{y|A_2y = v\}} \ \sup_{\{x|A_1x = u\}} \tilde{K}(x,y)$$

for $\tilde{K}$ any member of [K]. A mild hypothesis is given which ensures that indeed such a class exists and that all its members are closed. What is surprising is that this hypothesis is precisely the same as is needed to ensure the existence of the class $[K^*A^*]$ formed by the first operation from $[K^*]$ and $A^* = A_1^* \times A_2^*$. Furthermore, it is shown that under this hypothesis [AK] and $[K^*A^*]$ are conjugate classes. This is the analogue

of formula (1) for saddle-functions.

The development of these operations and the proof of the duality between them make up the main contribution of this thesis. Three forms of this duality are given. The most general version is proved in §1. In §2 a more explicit version is given; this is the form we find most useful for the subsequent applications. The formulation in §3 contains the sharpest conclusions and requires the strongest hypotheses.

In §4 the first application of this duality is made in defining addition and minimax convolution for saddle-functions and showing that these are dual operations. The result

$$\partial(K_1 + K_2)(x,y) = \partial K_1(x,y) + \partial K_2(x,y)$$

is also obtained for the subdifferential of the sum of two saddle-functions. This parallels the result for convex functions obtained by Rockafellar [38], Moreau [32], and others. The duality between addition and minimax convolution gives a general framework within which to consider problems of the form, "find the saddle-points of H with respect to $C \times D$, where H is a saddle-function and C and D are convex sets."

From the first application we obtain the following result. For $i = 1,\ldots,p$ let K_i be a closed concave-convex function on $R^m \times R^n$ which is not identically $+\infty$ or $-\infty$ and let T_i be the maximal monotone operator on R^{m+n} arising from the subdifferential of K_i (see [44, 46]). If each $R(T_i)$ is bounded, where $R(\cdot)$ denotes the range of an operator, then $\sum T_i$ is maximal monotone and

$$\sum \mathrm{cl}\ R(T_i) = \mathrm{cl}\ R(\sum T_i). \qquad (3)$$

It is known that this formula holds whenever the T_i's are subdifferentials of closed proper convex functions and each $R(T_i)$ is bounded. On the other hand, formula (3) fails in general for maximal monotone operators. However it

is not known whether formula (3) holds for arbitrary maximal monotone operators under the assumption that the sets $R(T_i)$ are bounded. But the fact that it holds for those maximal monotone operators arising from saddle-functions leads one to conjecture it holds in general. This is because such operators, unlike the subdifferentials of convex functions, exhibit most of the pathology of arbitrary maximal monotone operators. Indeed, the last fact is one of the main motivations for studying saddle-functions.

In §6 we make a second principal application of our fundamental dual operations in developing a perturbational duality theory for generalized saddle programs. We define a generalized saddle program to be an "objective" saddle-function K_0 (thought of as some given minimax problem) together with a particular class of perturbations. The entire program is given by another saddle-function K. To this generalized saddle program K we associate a dual generalized saddle program L. Under mild hypotheses on the perturbations in K, the dual program L has a unique (up to equivalence) "objective" saddle-function L_0. The minimax problem corresponding to L_0 is a dual to the original minimax problem. Optimal solutions, stable optimal solutions and Kuhn-Tucker vectors for these dual programs are studied. In §5, as a subsidiary application of the fundamental dual operations, we define a partial conjugacy correspondence among closed saddle-functions which is one-to-one and symmetric. By means of this correspondence we are able to associate with a generalized saddle program and its dual a well-defined Lagrangian saddle-function. We then give a characterization of the primal and dual stable optimal solutions and Kuhn-Tucker vectors in terms of the saddle-points of the Lagrangian.

In §7 this perturbational duality theory is used to study the problem of finding a saddle-point subject to finitely many convex and concave

constraints. Ordinary saddle programs are defined as a framework to treat such problems. A question of particular concern is whether or not a Lagrange multiplier principle holds for these saddle programs. The analogous question for ordinary convex programs (i.e. minimizing a convex function subject to finitely many convex constraints) has a very satisfying affirmative answer (see, for example, [44, Theorem 28.1]). It is shown that one cannot hope for a correspondingly general Lagrange multiplier principle for ordinary saddle programs. This is essentially due to the fact that, unlike the convex program case, the set of saddle-points of the Lagrangian does not split up into the product of the primal stable optimal solutions and the primal Kuhn-Tucker vectors (Lagrange multipliers). Put another way, the stable optimal solutions and Kuhn-Tucker vectors are shown to be in a certain sense dependent on each other.

Finally, in §8 the perturbational duality theory is specialized in another direction to yield a minimax version of Fenchel's Duality Theorem. We deal with dual pairs of minimax problems of the following form (where we suppress for simplicity now the issue of the domains of the variables):

(I) Find the saddle-points of $K(x,y) - LA(x,y)$.

(II) Find the saddle-points of $L^*(z,w) - K^*A^*(z,w)$.

Here K is closed and concave-convex on $R^m \times R^n$, L is closed and convex-concave on $R^p \times R^q$, and A is a (product) linear transformation from $R^m \times R^n$ to $R^p \times R^q$. The results obtained generalize certain results of Rockafellar [43], Lebedev-Tynjanskii [30], and Tynjanskii [57, 58].

It is known that many results in the theory of convex functions have refinements when polyhedralness is present. For closed saddle-functions there is a property of polyhedralness which is preserved under conjugacy as well as the operations in §§2, 4 and 5. Nearly all the results in the thesis

admit refinements when such polydedralness is present. This is discussed in the Appendix.

§0: Preliminaries

In this thesis we use mainly the definitions and notation set forth in Rockafellar [44], and any terms not defined in the thesis are to be understood as in [44]. For convenience we review some definitions here and also introduce some of our own. In addition, we present a few background results which will be of use later on.

The topology taken on R^n is the usual one, and the *interior* and *closure* of a subset S of R^n are denoted by $\text{int}\, S$ and $\text{cl}\, S$, respectively. A set is called *affine* iff it is either the empty set, denoted by $\emptyset$, or a translate of a linear subspace. The *affine hull* of a subset is the smallest affine set containing it. If C is a convex subset of R^n its *relative interior*, written $\text{ri}\, C$, is the interior of C with respect to its affine hull equipped with the relative topology.

If A is a linear transformation from R^p to R^m, then A^* denotes the *adjoint* linear transformation mapping R^m to R^p.

The *effective domain* of a convex function f on R^n is the set

$$\text{dom}\, f = \{x \mid f(x) < +\infty\},$$

and the *conjugate* of f is the convex function f^* on R^n given by

$$f^*(x^*) = \sup_x \{\langle x,x^*\rangle - f(x)\}$$

(where $\langle\cdot,\cdot\rangle$ denotes the ordinary inner product). Similarly, the effective domain of a concave function g on R^n is the set

$$\text{dom}\, g = \{x \mid g(x) > -\infty\},$$

and the conjugate of g is the concave function g^* on R^n given by

$$g^*(x^*) = \inf_x \{\langle x,x^*\rangle - g(x)\}.$$

Our multiple use of the superscript $*$ should cause no difficulty, since it is always clear from the context what operation is intended.

For any subset C of R^n the function $\delta(\cdot|C)$ on R^n, called the *indicator function* of C, is defined by setting $\delta(x|C)$ equal to 0 if

$x \in C$ and $+\infty$ otherwise. Clearly C is convex iff $\delta(\cdot|C)$ is convex, and in this case the conjugate of $\delta(\cdot|C)$ is denoted by $\delta^*(\cdot|C)$ and is given by

$$\delta^*(x^*|C) = \sup\{\langle x,x^*\rangle | x \in C\}.$$

We call $\delta^*(\cdot|C)$ the <u>support function</u> of C.

A <u>concave-convex</u> function on $R^m \times R^n$ is a function K from $R^m \times R^n$ to $[-\infty,+\infty]$ such that $K(x,y)$ is a concave function of $x \in R^m$ for each fixed $y \in R^n$ and a convex function of $y \in R^n$ for each fixed $x \in R^m$. A <u>convex-concave</u> function is defined the same except for interchanging "concave" with "convex." A <u>saddle-function</u> is either a concave-convex or a convex-concave function.

For the remainder of §0 let K denote a concave-convex function on $R^m \times R^n$.

We say that K has a saddle-value, or that the <u>saddle-value exists</u>, iff the two quantities

$$\sup_x \inf_y K(x,y)$$

and

$$\inf_y \sup_x K(x,y)$$

are equal, in which case this common value is the <u>saddle-value</u> of K. A pair $(\overline{x},\overline{y}) \in R^m \times R^n$ is a <u>saddle-point</u> of K iff

$$K(x,\overline{y}) \le K(\overline{x},\overline{y}) \le K(\overline{x},y)$$

for each $(x,y) \in R^m \times R^n$.

Define subsets $\mathrm{dom}_1 K$ of R^m and $\mathrm{dom}_2 K$ of R^n by

$$\mathrm{dom}_1 K = \{x | K(x,\cdot) \text{ is never } -\infty\},$$

$$\mathrm{dom}_2 K = \{y | K(\cdot,y) \text{ is never } +\infty\}.$$

The product set

$$\mathrm{dom}_1 K \times \mathrm{dom}_2 K = \mathrm{dom}\, K$$

is the _domain_ of K. We say that K is _proper_ iff its domain is nonempty. The _kernel_ of K is the restriction of K to the relative interior of its domain. We say that K is _simple_ iff $\operatorname{dom} K(x,\cdot) \subset \operatorname{cl}(\operatorname{dom}_2 K)$ for every $x \in \operatorname{ri}(\operatorname{dom}_1 K)$ and $\operatorname{dom} K(\cdot,y) \subset \operatorname{cl}(\operatorname{dom}_1 K)$ for every $y \in \operatorname{ri}(\operatorname{dom}_2 K)$. The function $\operatorname{cl}_1 K$ obtained by closing $K(x,y)$ as a concave function of x for each fixed y is called the _concave closure_ of K. Similarly, the function $\operatorname{cl}_2 K$ obtained by closing $K(x,y)$ as a convex function of y for each fixed x is called the _convex closure_ of K. If L is also a concave-convex function on $R^m \times R^n$, we say that K and L are _equivalent_ and write $K \sim L$ iff $\operatorname{cl}_1 K = \operatorname{cl}_1 L$ and $\operatorname{cl}_2 K = \operatorname{cl}_2 L$. The collection of all concave-convex functions on $R^m \times R^n$ which are equivalent to K is called the _equivalence class_ containing K and is denoted by $[K]$. We say that K is _closed_ iff $\operatorname{cl}_1 K \sim K$ and $K \sim \operatorname{cl}_2 K$.

It is an easy exercise to show that $K \sim L$ implies both that $\operatorname{dom} K = \operatorname{dom} L$ and that K is closed iff L is closed. Thus, an equivalence class is called closed (resp. proper) iff any of its members is closed (resp. proper).

The function f on $R^m \times R^n$ given by $f(x,y^*) = \sup_y \{\langle y,y^*\rangle - K(x,y)\}$ is convex in (x,y^*) jointly, since it is a pointwise supremum of convex functions. Similarly, the function g on $R^m \times R^n$ given by $g(x^*,y) = \inf_x \{\langle x,x^*\rangle - K(x,y)\}$ is concave in (x^*,y) jointly. We call f (resp. g) the _convex_ (resp. _concave_) _parent_ of K. Notice that this usage of the term "parent" differs by some minus signs from the original usage in Rockafellar [43]. With these definitions, (34.2) implies the following.

THEOREM 0.1. _Let_ f _(resp._ g_) be the convex (resp. concave) parent of_ K_. Then_ K _is closed iff_ $f(x,y^*) = -g^*(x,-y^*)$ _and_ $g(x^*,y) = -f^*(-x^*,y)$_, in which case_ (a) _and_ (b) _below hold._

(a) *For each* $\tilde{K} \in [K]$, *the convex* (*resp. concave*) *parent of* $\tilde{K}$ *is* f (*resp.* g). *Moreover*, f *and* g *are closed, and*

$$\mathrm{dom}_1 K = \{x \mid f(x,y^*) < +\infty \text{ for some } y^*\},$$

$$\mathrm{dom}_2 K = \{y \mid g(x^*,y) > -\infty \text{ for some } x^*\},$$

$$f(x,y^*) = -g^*(x,-y^*),$$

$$g(x^*,y) = -f^*(-x^*,y).$$

(b) *The equivalence class* $[K]$ *consists of all and only those concave-convex functions* $\tilde{K}$ *on* $R^m \times R^n$ *which satisfy* $\underline{K} \le \tilde{K} \le \overline{K}$, *where*

$$\underline{K}(x,y) = \sup_{y^*}\{\langle y^*,y\rangle - f(x,y^*)\}$$

and

$$\overline{K}(x,y) = \inf_{x^*}\{\langle x^*,x\rangle - g(x^*,y)\}.$$

Moreover, $\mathrm{cl}_2\tilde{K} = \underline{K}$ *and* $\mathrm{cl}_1\tilde{K} = \overline{K}$ *for each* $\tilde{K} \in [K]$, *and*

$$\underline{K}(x,y) = \overline{K}(x,y)$$

whenever $x \in \mathrm{ri}(\mathrm{dom}_1 K)$ *or* $y \in \mathrm{ri}(\mathrm{dom}_2 K)$.

The *lower conjugate* of K, denoted by $\underline{K}^*$, is a function on $R^m \times R^n$ defined by

$$\underline{K}^*(x^*,y^*) = \sup_y \inf_x \{\langle x,x^*\rangle + \langle y,y^*\rangle - K(x,y)\}.$$

Similarly, the *upper conjugate* of K, denoted by $\overline{K}^*$, is a function on $R^m \times R^n$ defined by

$$\overline{K}^*(x^*,y^*) = \inf_x \sup_y \{\langle x,x^*\rangle + \langle y,y^*\rangle - K(x,y)\}.$$

From (37.1) we have the following result.

THEOREM 0.2. *Assume* K *is closed. Then* $\underline{K}^*$ *and* $\overline{K}^*$ *are equivalent, closed concave-convex functions which depend only on* $[K]$. *Moreover, if* L *is any element of the equivalence class containing* $\underline{K}^*$ *and* $\overline{K}^*$, *then*

$$\mathrm{cl}_2 L = \underline{K}^*, \qquad \mathrm{cl}_1 L = \overline{K}^*,$$

$$\underline{L}^* = \underline{K}, \qquad \overline{L}^* = \overline{K},$$

and the convex (*resp. concave*) *parent of* L *is the negative of the concave*

(*resp. convex*) *parent of* K.

The equivalence class containing $\underline{K}^*$ and $\overline{K}^*$ is called *the conjugate* of $[K]$ and is denoted by $[K^*]$. Each member of $[K^*]$ is called *a conjugate* of every member of $[K]$. It is immediate from Theorems 0.2 and 0.1(b) that (at least when K is closed) $\underline{K}^*$ and $\overline{K}^*$ are the least and greatest elements of $[K^*]$, respectively. The notation thus conforms to that introduced in Theorem 0.1(b), where a lower (resp. upper) bar indicates the least (resp. greatest) element of the equivalence class.

By (34.2.3) the only equivalence classes which are closed but not proper are the one containing the constant function $+\infty$ and the one containing the constant function $-\infty$. Since each of these two equivalence classes is the conjugate of the other, it follows that $[K^*]$ is closed and proper whenever $[K]$ is.

Suppose for now that K is closed and proper, and let K^* be a conjugate of K. By (37.1.3) the saddle-value of K exists whenever $0 \in \operatorname{ri}(\operatorname{dom}_1 K^*)$ or $0 \in \operatorname{ri}(\operatorname{dom}_2 K^*)$, and by (37.5.3) a saddle-point of K exists when both these conditions are satisfied. To use these facts it would be helpful to have characterizations of $0 \in \operatorname{ri}(\operatorname{dom}_j K^*)$ for $j = 1$ and 2. These are furnished by the next two lemmas.

LEMMA 0.3. *Assume* K *is closed, and let* f (*resp.* g) *be its convex* (*resp. concave*) *parent. Then*

$$\operatorname{ri}(\operatorname{dom}_1 K^*) = \bigcup\{\operatorname{ri}(\operatorname{dom} g(\cdot,y)) \mid y \in \operatorname{ri} D\}$$

and

$$\operatorname{ri}(\operatorname{dom}_2 K^*) = \bigcup\{\operatorname{ri}(\operatorname{dom} f(x,\cdot)) \mid x \in \operatorname{ri} C\},$$

where $C = \operatorname{dom}_1 K$ *and* $D = \operatorname{dom}_2 K$. *These formulas also hold when "ri" is deleted throughout.*

PROOF. From Theorems 0.2 and 0.1 (a) it follows that

$$\mathrm{dom}_1 K^* = A \,\mathrm{dom}\, g$$

and

$$\mathrm{dom}\, g = \bigcup\{\mathrm{dom}\, g(\cdot,y) \times \{y\} \mid y \in D\},$$

where A is the projection $(x^*,y) \to x^*$. Hence (6.6) implies

$$\mathrm{ri}(\mathrm{dom}_1 K^*) = A\, \mathrm{ri}(\mathrm{dom}\, g)$$

and (6.8) implies

$$\mathrm{ri}(\mathrm{dom}\, g) = \bigcup\{\mathrm{ri}(\mathrm{dom}\, g(\cdot,y)) \times \{y\} \mid y \in \mathrm{ri}\, D\}.$$

The formulas for $\mathrm{dom}_1 K^*$ and its relative interior follow from these, and the other two formulas are proved similarly.

For the next lemma more definitions are needed. If f is a proper convex function on R^n, the <u>recession function</u> of f, written $\mathrm{rec}\, f$, is a function on R^n defined by

$$(\mathrm{rec}\, f)(y) = \sup\{f(x+y) - f(x) \mid x \in \mathrm{dom}\, f\},$$

and the <u>recession cone</u> of f is the set

$$\mathrm{rec\ cone}\, f = \{y \mid (\mathrm{rec}\, f)(y) \leq 0\}.$$

The recession function and recession cone of a proper concave function are defined similarly by replacing "sup" by "inf" and "$\leq$" by "$\geq$". This notation for these objects differs from that in [44].

Now write $C = \mathrm{dom}_1 K$ and $D = \mathrm{dom}_2 K$. The <u>convex recession function</u> of K is the function $\mathrm{rec}_2 K$ on R^n defined by

$$(\mathrm{rec}_2 K)(w) = \sup\{(\mathrm{rec}\, K(x,\cdot))(w) \mid x \in \mathrm{ri}\, C\}.$$

The <u>convex recession cone</u> of K is the set

$$\mathrm{rec\ cone}_2 K = \{w \mid (\mathrm{rec}_2 K)(w) \leq 0\}.$$

Similarly, the <u>concave recession function</u> of K is the function $\mathrm{rec}_1 K$ on R^m defined by

$$(\mathrm{rec}_1 K)(z) = \inf\{(\mathrm{rec}\, K(\cdot,y))(z) \mid y \in \mathrm{ri}\, D\},$$

and the <u>concave recession cone</u> of K is the set

$$\text{rec cone}_1 K = \{z \mid (\text{rec}_1 K)(z) \geq 0\}.$$

Trivially,

$$\text{rec cone}_2 K = \bigcap \{\text{rec cone } K(x,\cdot) \mid x \in \text{ri } C\}$$

and

$$\text{rec cone}_1 K = \bigcap \{\text{rec cone } K(\cdot,y) \mid y \in \text{ri } D\}.$$

When K is closed and proper, Theorem 0.1 (b) implies that the recession functions and cones of K depend only on $[K]$, and in fact, (37.2) states that

$$\text{rec}_2 K = \delta^*(\cdot \mid \text{dom}_2 K^*)$$

and

$$\text{rec}_1 K = -\delta^*(-\cdot \mid \text{dom}_1 K^*).$$

Furthermore, when K is closed and proper it follows from (34.3) and (8.5) that the recession cones of K are closed convex cones containing the origin. Hence they are subspaces iff they are closed under multiplication by -1.

LEMMA 0.4. <u>Assume</u> K <u>is closed and proper</u>. <u>Let</u> j <u>equal</u> 1 <u>or</u> 2 <u>and put</u> $S_j = \text{rec cone}_j K$. <u>Then</u>

$$0 \in \text{ri}(\text{dom}_j K^*) \ \underline{\text{iff}} \ S_j \subset -S_j,$$

<u>and</u>

$$0 \in \text{int}(\text{dom}_j K^*) \ \underline{\text{iff}} \ S_j \subset \{0\}.$$

PROOF. We use the following SUBLEMMA. If $w^* \in R^n$ and h is a positively homogeneous proper convex function on R^n, then the following two conditions are equivalent:

(a) $\forall w \in R^n$, $\langle w,w^*\rangle \leq h(w)$ with strict inequality for each w such that $-h(-w) \neq h(w)$;

(b) $\forall w \in R^n$, $h(w) \leq \langle w,w^*\rangle$ implies $h(-w) \leq \langle -w,w^*\rangle$.

PROOF OF SUBLEMMA. Assume (a) and suppose $h(w) \leq \langle w,w^*\rangle$. Then actually

$h(w) = \langle w,w^*\rangle$. If we had $-h(-w) \neq h(w)$, then (a) would imply $\langle w,w^*\rangle < h(w)$, a contradiction. Thus $-h(-w) = h(w) = \langle w,w^*\rangle$ and (b) is proved. Conversely, assume (b) and let w be given. If $h(w) \leq \langle w,w^*\rangle$, then (4.7.2) and (b) imply $-h(w) \leq h(-w) \leq \langle -w,w^*\rangle$ and hence $\langle w,w^*\rangle \leq h(w)$. Suppose $-h(-w) \neq h(w)$. By (4.7.2) we have $-h(-w) < h(w)$. If $h(w) \leq \langle w,w^*\rangle$, this would imply $-h(-w) < \langle w,w^*\rangle$ while at the same time from (b) we would have $\langle w,w^*\rangle \leq -h(-w)$. Therefore $\langle w,w^*\rangle < h(w)$ whenever $-h(-w) \neq h(w)$, and (a) is proved.

Define $h = \delta^*(\cdot|\mathrm{dom}_2K^*)$. By (13.1) and the Sublemma, $0 \in \mathrm{ri}(\mathrm{dom}_2K^*)$ iff for each $w \in R^n$, $h(w) \leq 0$ implies $h(-w) \leq 0$. By (13.1) we also have that $0 \in \mathrm{int}(\mathrm{dom}_2K^*)$ iff for each $w \in R^n$, $h(w) \leq 0$ implies $w = 0$. The assertions for $j = 2$ follow from these equivalences and the fact that $h = \mathrm{rec}_2K$. The assertions for $j = 1$ follow similarly, using the fact that $\mathrm{rec}_1K = -\delta^*(-\cdot|\mathrm{dom}_1K^*)$.

In view of Lemmas 0.3 and 0.4 and the observations preceding Lemma 0.3 concerning the existence of saddle-values and saddle-points, formulas are given in §§2 and 4 for the parents and recession functions of the saddle-functions resulting from the operations developed there. This is done in Theorem 2.4, Corollary 4.6.2, and Lemmas 2.5 and 4.7. By combining these results with Lemmas 0.3 and 0.4 the reader can easily state existence theorems as needed.

The next lemma is used in §§2 and 4 to dualize various conditions. Also, taking $L_j = \{0\}$ in this lemma yields the assertion "$0 \in \mathrm{ri}(\mathrm{dom}_jK^*)$ iff $S_j \subset -S_j$" of Lemma 0.4.

LEMMA 0.5. _Let_ L_1 _(resp._ L_2_) be a subspace of_ R^m _(resp._ R^n_), and assume_ K _is closed and proper. Then for_ $j = 1$ _and_ 2 _the following conditions are equivalent:_

(a_j) $L_j \cap \mathrm{ri}(\mathrm{dom}_j K^*) \neq \emptyset$;

(b_j) $L_j \cap (\text{rec cone}_j K)$ <u>is a subspace</u>;

(c_j) $L_j \cap (\text{rec cone}_j K) \subset -(\text{rec cone}_j K)$.

PROOF. We prove only that (a_2), (b_2) and (c_2) are equivalent, as the proof for $j = 1$ is similar. By the remarks preceding Lemma 0.4 and the fact that L_2 is a subspace, (b_2) is equivalent to (c_2). Write $D^* = \mathrm{dom}_2 K^*$ and $L = L_2$. By (11.3), (a_2) fails iff there exists a hyperplane separating L and D^*. By (11.1) this occurs iff there exists a $w \in R^n$ such that

$$\inf\{\langle y^*,w\rangle \mid y^* \in L\} \geq \sup\{\langle y^*,w\rangle \mid y^* \in D^*\}$$

and

$$\sup\{\langle y^*,w\rangle \mid y^* \in L\} > \inf\{\langle y^*,w\rangle \mid y^* \in D^*\}.$$

But since

$$\sup\{\langle y^*,w\rangle \mid y^* \in D^*\} = (\mathrm{rec}_2 K)(w)$$

and

$$\inf\{\langle y^*,w\rangle \mid y^* \in L\} = \begin{cases} 0 & \text{if } w \in L \\ -\infty & \text{if } w \notin L, \end{cases}$$

this means that (a_2) fails iff there exists a $w \in L$ such that $(\mathrm{rec}_2 K)(w) \leq 0$ and $(\mathrm{rec}_2 K)(-w) > 0$. Therefore ($a_2$) holds iff for each $w \in L^\perp$, $(\mathrm{rec}_2 K)(w) \leq 0$ implies $(\mathrm{rec}_2 K)(-w) \leq 0$. But this last condition is the same as (c_2).

Define K to be <u>polyhedral</u> iff it is closed and either its convex or its concave parent is polyhedral. By Theorem 0.1, if K is polyhedral and L is equivalent to K, then L is polyhedral. Thus, an equivalence class is called polyhedral iff any of its members is polyhedral.

From Theorem 0.2 follows the important fact that $[K^*]$ is polyhedral whenever $[K]$ is polyhedral. Polyhedralness is also preserved by each of the operations developed in §§2, 4 and 5.

§1: A General Theorem

The goal of this section is to prove Theorem 1.1, which concerns the following question. Let K be a saddle-function on $R^m \times R^n$, let $A_1: R^p \to R^m$ and $A_2: R^q \to R^n$ be linear transformations, and suppose that all the saddle-functions of the form $(u,v) \to \tilde{K}(A_1u, A_2v)$, where $\tilde{K}$ is equivalent to K, belong to a single equivalence class. What can be said about the conjugate (i.e. "dual") equivalence class? Theorem 1.1 describes this class explicitly. In §2 this duality is developed in more detail as two dual operations on equivalence classes.

THEOREM 1.1. *Let* K *be a concave-convex function on* $R^m \times R^n$ *and let* $A = A_1 \times A_2$, *where* $A_1: R^p \to R^m$ *and* $A_2: R^q \to R^n$ *are linear transformations. Assume* K *is closed and that there exists a closed concave-convex function* H *on* $R^p \times R^q$ *such that* $\tilde{K}A$ *is equivalent to* H *whenever* $\tilde{K}$ *is equivalent to* K. *Define* J_1 *and* J_2 *on* $R^p \times R^q$ *by*

$$J_1(u^*,v^*) = \sup_{\{x^* | A_1^* x^* = u^*\}} \ \inf_{\{y^* | A_2^* y^* = v^*\}} \underline{K}^*(x^*,y^*)$$

and

$$J_2(u^*,v^*) = \inf_{\{y^* | A_2^* y^* = v^*\}} \ \sup_{\{x^* | A_1^* x^* = u^*\}} \overline{K}^*(x^*,y^*),$$

and let J *be any concave-convex function on* $R^p \times R^q$ *such that* $J_1 \le J \le J_2$. *Then* J *is simple and satisfies*

$$cl_2 cl_1 J = \underline{H}^*, \quad cl_1 cl_2 J = \overline{H}^*,$$
$$\mathrm{dom}(cl_2 cl_1 J) = \mathrm{dom}\, H^* = \mathrm{dom}(cl_1 cl_2 J),$$

where

$$\mathrm{dom}\, H^* \subset cl(A^* \mathrm{dom}\, K^*).$$

If H *is proper, then* J *is proper and has the same kernel as* H^*.

From Theorem 0.2 and Lemma 1.6 (below) it follows that in the theorem J can be taken to be either J_1 or J_2.

The proof of this theorem uses seven lemmas. The assertion of Lemma 1.2

was first noted in Rockafellar [39] and the proof given is the one indicated there. Lemma 1.3 was suggested by the proof of (34.5). Lemma 1.4 was first proved in [39].

LEMMA 1.2. Let K be a concave-convex function. Then the lower conjugate of K^* is $cl_2 cl_1 K$, and the upper conjugate of $\bar{K}^*$ is $cl_1 cl_2 K$.

PROOF. Observe that for any convex function f,

$$(cl\ f)(y) = \sup_{y^*} \inf_{w} \{\langle y - w, y^*\rangle + f(w)\}$$

follows trivially from the fact that $cl\ f = (f^*)^*$. Similarly,

$$(cl\ g)(x) = \inf_{x^*} \sup_{z} \{\langle x - z, x^*\rangle + g(z)\}$$

for any concave function g. If H denotes the upper conjugate of K^*, then these facts together with the definitions imply

$$H(x,y) = \inf_{x^*} \sup_{y^*} \{\langle x^*,x\rangle + \langle y^*,y\rangle - \inf_{z} \sup_{w} \{\langle z,x^*\rangle + \langle w,y^*\rangle - K(z,w)\}\}$$

$$= \inf_{x^*} \sup_{z} \{\langle x - z,x^*\rangle + \sup_{y^*} \inf_{w} \{\langle y - w,y^*\rangle + K(z,w)\}\}$$

$$= \inf_{x^*} \sup_{z} \{\langle x - z,x^*\rangle + (cl_2 K)(z,y)\}$$

$$= (cl_1 cl_2 K)(x,y).$$

The other assertion is proved similarly.

LEMMA 1.3. Let K be a concave-convex function. Then $\mathrm{dom}_1 K = \mathrm{dom}_1(cl_2 K)$, $\mathrm{dom}_2 K \subset \mathrm{dom}_2(cl_2 K)$, and $\mathrm{dom}\ cl_2 K(\cdot,y) = \mathrm{dom}_1 K$ for every y. If $\mathrm{dom}_1 K \neq \emptyset$ and $\mathrm{dom}\ K(x,\cdot) \subset cl(\mathrm{dom}_2 K)$ for every $x \in ri(\mathrm{dom}_1 K)$, then actually $\mathrm{dom}_2 K \subset \mathrm{dom}_2(cl_2 K) \subset cl(\mathrm{dom}_2 K)$ and moreover $cl_2 K$ agrees with K on the set $ri(\mathrm{dom}\ K)$. Parallel assertions hold concerning $cl_1 K$.

PROOF. By the definition of $\mathrm{dom}_1 K$ and $\mathrm{dom}_2 K$, when $x \notin \mathrm{dom}_1 K$ the convex function $K(x,\cdot)$ somewhere has the value $-\infty$, whereas when $x \in \mathrm{dom}_1 K$ the function $K(x,\cdot)$ is never $-\infty$ and its effective domain includes $\mathrm{dom}_2 K$. Thus $(cl_2 K)(x,\cdot)$ is the constant function $-\infty$ when $x \notin \mathrm{dom}_1 K$, whereas

when $x \in \mathrm{dom}_1 K$ it is never $-\infty$ and its effective domain includes $\mathrm{dom}_2 K$. This shows that $\mathrm{dom}_1(\mathrm{cl}_2 K) = \mathrm{dom}_1 K$ and $\mathrm{dom}_2(\mathrm{cl}_2 K) \supset \mathrm{dom}_2 K$, and in fact that $\mathrm{dom}_1 K$ is the effective domain of every one of the concave functions $(\mathrm{cl}_2 K)(\cdot,y)$. Now assume that $\mathrm{dom}\, K(x,\cdot) \subset \mathrm{cl}(\mathrm{dom}_2 K)$ for every $x \in \mathrm{ri}(\mathrm{dom}_1 K)$. Since $\mathrm{dom}_2 K \subset \mathrm{dom}\, K(x,\cdot)$ always holds, (6.3.1) implies

$$\mathrm{ri}(\mathrm{dom}\, K(x,\cdot)) = \mathrm{ri}(\mathrm{dom}_2 K) \tag{1}$$

for every $x \in \mathrm{ri}(\mathrm{dom}_1 K)$. Thus, for each $x \in \mathrm{ri}(\mathrm{dom}_1 K)$ the convex function $K(x,\cdot)$ agrees with its closure on $\mathrm{ri}(\mathrm{dom}_2 K)$. That is, K agrees with $\mathrm{cl}_2 K$ on the set $\mathrm{ri}(\mathrm{dom}_1 K) \times \mathrm{ri}(\mathrm{dom}_2 K) = \mathrm{ri}(\mathrm{dom}\, K)$. Finally, assume $\mathrm{dom}_1 K \neq \emptyset$. By (6.2) we can pick some $x \in \mathrm{ri}(\mathrm{dom}_1 K)$. Then $x \in \mathrm{dom}_1(\mathrm{cl}_2 K)$ implies $\mathrm{dom}(\mathrm{cl}_2 K)(x,\cdot) \subset \mathrm{cl}(\mathrm{dom}\, K(x,\cdot))$, and equation (1) and (6.3.1) imply $\mathrm{cl}(\mathrm{dom}\, K(x,\cdot)) = \mathrm{cl}(\mathrm{dom}_2 K)$. Thus, for any $y \in \mathrm{dom}_2(\mathrm{cl}_2 K)$, $(\mathrm{cl}_2 K)(x,y) < +\infty$ implies that $y \in \mathrm{dom}(\mathrm{cl}_2 K)(x,\cdot)$ and hence $y \in \mathrm{cl}(\mathrm{dom}_2 K)$. This shows $\mathrm{dom}_2(\mathrm{cl}_2 K) \subset \mathrm{cl}(\mathrm{dom}_2 K)$.

LEMMA 1.4. <u>Let</u> K <u>be a concave-convex function</u>. <u>Then</u> $\mathrm{cl}_2 K$ <u>and</u> $\mathrm{cl}_1 K$ <u>are simple</u>.

PROOF. It suffices to prove $\mathrm{cl}_2 K$ is simple. This requires showing (i) $\mathrm{dom}(\mathrm{cl}_2 K)(\cdot,y) \subset \mathrm{cl}(\mathrm{dom}_1(\mathrm{cl}_2 K))$ whenever $y \in \mathrm{ri}(\mathrm{dom}_2(\mathrm{cl}_2 K))$, and (ii) $\mathrm{dom}(\mathrm{cl}_2 K)(x,\cdot) \subset \mathrm{cl}(\mathrm{dom}_2(\mathrm{cl}_2 K))$ whenever $x \in \mathrm{ri}(\mathrm{dom}_1(\mathrm{cl}_2 K))$. Lemma 1.3 implies $\mathrm{dom}_1(\mathrm{cl}_2 K) = \mathrm{dom}_1 K$ and $\mathrm{dom}(\mathrm{cl}_2 K)(\cdot,y) = \mathrm{dom}_1 K$ for every y. This establishes (i). Since a concave function agrees with its closure on the relative interior of its effective domain (see (7.2) and (7.4)), Lemma 1.3 also implies $(\mathrm{cl}_2 K)(x,y) = \mathrm{cl}((\mathrm{cl}_2 K)(\cdot,y))(x) = (\mathrm{cl}_1 \mathrm{cl}_2 K)(x,y)$ whenever $x \in \mathrm{ri}(\mathrm{dom}_1(\mathrm{cl}_2 K))$. Applying Lemma 1.3 once more yields $\mathrm{dom}(\mathrm{cl}_1 \mathrm{cl}_2 K)(x,\cdot) = \mathrm{dom}_2(\mathrm{cl}_2 K)$ for every x. Combining these facts establishes (ii).

LEMMA 1.5. <u>Let</u> K <u>be a concave-convex function</u>. <u>If</u> $\mathrm{ri}(\mathrm{dom}(\mathrm{cl}_2 \mathrm{cl}_1 K)) = \mathrm{ri}(\mathrm{dom}(\mathrm{cl}_1 \mathrm{cl}_2 K))$, <u>then</u> K <u>is simple</u>.

PROOF. Assume there exists an $x \in \mathrm{ri}(\mathrm{dom}_1 K)$ such that $\mathrm{dom}\, K(x,\cdot) \not\subset \mathrm{cl}(\mathrm{dom}_2 K)$. Then there exists a y such that $K(x,y) < +\infty$ but $y \notin \mathrm{cl}(\mathrm{dom}_2 K)$. Lemma 1.3 implies $\mathrm{dom}(\mathrm{cl}_2 K)(\cdot,y) = \mathrm{dom}_1 K$. Since an improper concave function euqals $+\infty$ everywhere on the relative interior of its effective domain (by(7.2)), $(\mathrm{cl}_2 K)(x,y) \leq K(x,y) < +\infty$ implies that $(\mathrm{cl}_2 K)(\cdot,y)$ is never $+\infty$. Hence $y \in \mathrm{dom}_2(\mathrm{cl}_2 K)$. From Lemma 1.3 it follows that

$$\mathrm{dom}_2 K \subset \mathrm{dom}_2(\mathrm{cl}_2 K) = \mathrm{dom}_2(\mathrm{cl}_1 \mathrm{cl}_2 K) \not\subset \mathrm{cl}(\mathrm{dom}_2 K).$$

By (6.3.1) this implies

$$\mathrm{ri}(\mathrm{dom}_2(\mathrm{cl}_1 \mathrm{cl}_2 K)) \neq \mathrm{ri}(\mathrm{dom}_2 K). \tag{1}$$

By (33.1.1) $\mathrm{cl}_1 K$ is a concave-convex function. Hence Lemma 1.3 implies

$$\mathrm{dom}_2 K = \mathrm{dom}_2(\mathrm{cl}_1 K) \subset \mathrm{dom}_2(\mathrm{cl}_2 \mathrm{cl}_1 K)$$

and $\mathrm{dom}_1 K \subset \mathrm{dom}_1(\mathrm{cl}_1 K)$. Since $x \in \mathrm{ri}(\mathrm{dom}_1 K)$, this shows $\mathrm{dom}_1(\mathrm{cl}_1 K) \neq \emptyset$, and by Lemma 1.4 $\mathrm{cl}_1 K$ is simple. Hence Lemma 1.3 also implies

$$\mathrm{dom}_2(\mathrm{cl}_2 \mathrm{cl}_1 K) \subset \mathrm{cl}(\mathrm{dom}_2(\mathrm{cl}_1 K)).$$

These facts together with (6.3.1) imply that

$$\mathrm{ri}(\mathrm{dom}_2(\mathrm{cl}_2 \mathrm{cl}_1 K)) = \mathrm{ri}(\mathrm{dom}_2 K). \tag{2}$$

From (1) and (2) it follows that $\mathrm{ri}(\mathrm{dom}_2(\mathrm{cl}_2 \mathrm{cl}_1 K)) = \mathrm{ri}(\mathrm{dom}_2(\mathrm{cl}_1 \mathrm{cl}_2 K))$ implies $\mathrm{dom}\, K(x,\cdot) \subset \mathrm{cl}(\mathrm{dom}_2 K)$ for every $x \in \mathrm{ri}(\mathrm{dom}_1 K)$. The other condition for K to be simple can be established similarly.

LEMMA 1.6. *Let* L *be a concave-convex function on* $R^m \times R^n$ *and* $B = B_1 \times B_2$ *be a linear transformation from* $R^m \times R^n$ *to* $R^p \times R^q$. *Define*

$$J_1(z,w) = \sup_{\{x|B_1x = z\}} \inf_{\{y|B_2y = w\}} L(x,y)$$

and

$$J_2(z,w) = \inf_{\{y|B_2y = w\}} \sup_{\{x|B_1x = z\}} L(x,y)$$

for every $(z,w) \in R^p \times R^q$. *Then* J_1 *and* J_2 *are concave-convex functions on* $R^p \times R^q$.

PROOF. Let $z \in R^p$ be fixed and write $f = J_1(z,\cdot)$. Then $f(w) = \sup\{(B_2L(x,\cdot))(w) \mid B_1x = z\}$. Since (5.7) implies $B_2L(x,\cdot)$ is a convex function for each x, (5.5) implies f itself is convex. Now let $w \in R^q$ be fixed and write $g = J_1(\cdot,w)$. The function $k(x) = \inf\{L(x,y) \mid B_2y = w\}$ is concave by (5.5), so $g = B_1k$ is concave by (5.7). This shows J_1 is concave-convex. Similarly, (5.5) and (5.7) imply J_2 is concave-convex.

LEMMA 1.7. <u>Let the notation be as in Lemma 1.6. Then</u> $\underline{L}^*B^* \le \underline{J}_2^*$ <u>and</u> $\overline{J}_1^* \le \overline{L}^*B^*$, <u>where</u> $B^* = B_1^* \times B_2^*$.

PROOF. For $j = 1$ and 2 let R_j denote the range of B_j. Observe that $J_2(\cdot,w)$ is constantly $+\infty$ whenever $w \notin R_2$, and $J_2(z,w) = -\infty$ whenever $w \in R_2$ but $z \notin R_1$. These facts and the definition of lower conjugate yield

$$\begin{aligned}
\underline{J}_2^*(z^*,w^*) &= \sup_w \inf_z \{\langle z,z^*\rangle + \langle w,w^*\rangle - J_2(z,w)\} \\
&= \sup_{w\in R_2} \inf_{z\in R_1} \{\langle z,z^*\rangle + \langle w,w^*\rangle - J_2(z,w)\} \\
&= \sup_{w\in R_2} \inf_{z\in R_1} \sup_{y\in B_2^{-1}w} \inf_{x\in B_1^{-1}z} \{\langle z,z^*\rangle + \langle w,w^*\rangle - L(x,y)\} \\
&\ge \sup_{w\in R_2} \sup_{y\in B_2^{-1}w} \inf_{z\in R_1} \inf_{x\in B_1^{-1}z} \{\langle z,z^*\rangle + \langle w,w^*\rangle - L(x,y)\} \\
&= \sup_{\bar{y}} \sup_{y\in\bar{y}+B_2^{-1}0} \inf_{\bar{x}} \inf_{x\in\bar{x}+B_1^{-1}0} \{\langle B_1\bar{x},z^*\rangle + \langle B_2\bar{y},w^*\rangle - L(x,y)\} \\
&= \sup_y \inf_x \{\langle x,B_1^*z^*\rangle + \langle y,B_2^*w^*\rangle - L(x,y)\} \\
&= \underline{L}^*(B_1^*z^*,B_2^*w^*) \\
&= \underline{L}^*B^*(z^*,w^*),
\end{aligned}$$

where the inequality follows from (36.1). This proves the first inequality, and the second is proved similarly.

LEMMA 1.8. <u>Let</u> K <u>and</u> A <u>be as in Theorem 1.1, write</u> $C = \mathrm{dom}_1K$ <u>and</u> $D = \mathrm{dom}_2K$, <u>and let</u> f <u>(resp.</u> g<u>) be the convex (resp. concave) parent of</u> K.

If K *is concave-closed, then*

$$\mathrm{dom}_1(\underline{KA})^* = \bigcup\{\mathrm{dom}\ \mathrm{cl}(A_1^* g(\cdot,y)) \mid y \in D \cap \text{range } A_2\}.$$

If K *is convex-closed, then*

$$\mathrm{dom}_2(\overline{KA})^* = \bigcup\{\mathrm{dom}\ \mathrm{cl}(A_2^* f(x,\cdot)) \mid x \in C \cap \text{range } A_1\}.$$

PROOF. We prove only the first formula, as the second is similar. Assume K is concave-closed. Then (16.3) implies $(K(\cdot,y)A_2)^* = \mathrm{cl}(A_2^* g(\cdot,y))$ for every y. Also, if $y \notin D$ then $K(\cdot,y)$ is $+\infty$ somewhere, which implies that $g(\cdot,y) = K(\cdot,y)^*$ and hence $\mathrm{cl}(A_2^* g(\cdot,y))$ are constantly $-\infty$. From these facts it follows that

$$\begin{aligned}(\underline{KA})^*(u^*,v^*) &= \sup_v \inf_u \{\langle u,u^*\rangle + \langle v,v^*\rangle - K(A_1u,A_2v)\} \\ &= \sup_v \{\langle v,v^*\rangle + (K(\cdot,A_2v)A_1)^*(u^*)\} \\ &= \sup_{v \in A_2^{-1}D} \{\langle v,v^*\rangle + \mathrm{cl}(A_1^* g(\cdot,A_2v))(u^*)\}.\end{aligned}$$

Hence $u^* \in \mathrm{dom}_1(\underline{KA})^*$ iff

$$\forall v^*,\ \exists v \in A_2^{-1}D \text{ such that } u^* \in \mathrm{dom}\ \mathrm{cl}(A_1^* g(\cdot,A_2v)),$$

and this occurs iff

$$\exists y \in D \cap \text{range } A_2 \text{ such that } u^* \in \mathrm{dom}\ \mathrm{cl}(A_1^*, g(\cdot,y)).$$

The first formula is immediate from this.

PROOF OF THEOREM 1.1. Since K is closed, Theorem 0.2 implies that the lower conjugate of $\overline{K}^*$ is $\underline{K}$ and the upper conjugate of $\underline{K}^*$ is $\overline{K}$. Thus, Lemma 1.7 implies $\underline{KA} \le \underline{J}^{\ddagger}$ and $\overline{J}^{\dagger} \le \overline{KA}$. From $J_1 \le J \le J_2$ and (36.1) it follows that $\underline{J}^{\ddagger} \le \underline{J}^* \le \overline{J}^* \le \overline{J}^{\dagger}$. Therefore

$$\underline{KA} \le \underline{J}^* \le \overline{J}^* \le \overline{KA}.$$

Together with Theorems 0.1 and 0.2, this implies that the lower conjugates of $\underline{J}^*$ and H coincide and the upper conjugates of $\overline{J}^*$ and H coincide. But by Lemma 1.2 this means that

$$\mathrm{cl}_2\mathrm{cl}_1 J = \underline{H}^* \quad\text{and}\quad \mathrm{cl}_1\mathrm{cl}_2 J = \overline{H}^*.$$

Since equivalent saddle-functions have the same domain, this implies that

$$\mathrm{dom}(\mathrm{cl}_2\mathrm{cl}_1 J) = \mathrm{dom}\, H^* = \mathrm{dom}(\mathrm{cl}_1\mathrm{cl}_2 J) \tag{1}$$

By the hypothesis on H and Theorems 0.1 and 0.2 we have that

$$\mathrm{dom}\, H^* = \mathrm{dom}_1(\underline{K}A)^* \times \mathrm{dom}_2(\overline{K}A)^*. \tag{2}$$

Let f, g, C and D be as in Lemma 1.8. Since K is closed, Theorem 0.1 implies that $\overline{K}$ is concave-closed, $\underline{K}$ is convex-closed, and both $\overline{K}$ and $\underline{K}$ have the same parents as K. Hence Lemma 1.8 implies, for example, that

$$\mathrm{dom}_2(\overline{K}A)^* = \bigcup\{\mathrm{dom}\ \mathrm{cl}(A_2^* f(x,\cdot)) \mid x \in C \cap \mathrm{range}\ A_2\}.$$

Since

$$\mathrm{dom}\ \mathrm{cl}(A_2^* f(x,\cdot)) \subset \mathrm{cl}(\mathrm{dom}\ A_2^* f(x,\cdot)) = \mathrm{cl}(A_2^* \mathrm{dom}\ f(x,\cdot))$$

for each x, we conclude that

$$\mathrm{dom}_2(\overline{K}A)^* \subset \mathrm{cl}(A_2^* \bigcup\{\mathrm{dom}\ f(x,\cdot) \mid x \in C \cap \mathrm{range}\ A_1\}).$$

But by Lemma 0.3 we know that

$$\mathrm{dom}_2 K^* = \bigcup\{\mathrm{dom}\ f(x,\cdot) \mid x \in C\}.$$

Using (2), it follows that

$$\mathrm{dom}_j H^* \subset \mathrm{cl}(A_j^* \mathrm{dom}_j K^*)$$

for $j = 2$. The corresponding inclusion for $j = 1$ follows from (2) similarly. Hence

$$\mathrm{dom}\, H^* \subset \mathrm{cl}(A^* \mathrm{dom}\, K^*).$$

From (1) and Lemma 1.5 it follows that J is simple. Now assume H is proper. Then the set in (1) is nonempty. If $\mathrm{dom}_1 J$ were empty, then $\mathrm{cl}_2 J$ and hence $\mathrm{cl}_1\mathrm{cl}_2 J$ would be constantly $-\infty$, contradicting the fact that $\mathrm{dom}_1(\mathrm{cl}_1\mathrm{cl}_2 J)$ is nonempty. Hence $\mathrm{dom}_1 J \neq \emptyset$. A similar argument shows $\mathrm{dom}_2 J \neq \emptyset$, so that J is proper. Since J is also simple, it follows from (6.3.1) and Lemmas 1.3 and 1.4 that J has the same kernel as $\mathrm{cl}_1 J$, which in turn has the same kernel as $\mathrm{cl}_2\mathrm{cl}_1 J$. Then since $\mathrm{cl}_2\mathrm{cl}_1 J \sim H^*$, Theorem 0.1 (b) implies that J has the same kernel as H^*.

§2: Two Dual Operations

In this section we develop two fundamental operations involving linear transformations and equivalence classes of closed proper saddle-functions. Specific conditions are given under which the operations can be performed, and the operations are shown to be dual. Various results are proved concerning the equivalence classes resulting from these operations. The section concludes with examples showing that the conditions under which the operations can be performed cannot in general be weakened.

The first operation we develop is analogous to that of composing a convex function with a linear transformation. Let K be a closed proper concave-convex function on $R^m \times R^n$ and let $A = A_1 \times A_2$ be a linear transformation from $R^p \times R^q$ to $R^m \times R^n$. We seek a condition on K and A ensuring the existence of an equivalence class of closed proper saddle-functions which contains every function of the form $\tilde{K}A$ for $\tilde{K} \in [K]$. Such a condition is given in Theorem 2.2. When this equivalence class exists, it will usually be denoted by $[KA]$.

LEMMA 2.1. *Let K be a concave-convex function on $R^m \times R^n$ and let $A = A_1 \times A_2$ be a linear transformation from $R^p \times R^q$ to $R^m \times R^n$. Then KA is a concave-convex function on $R^p \times R^q$, and $A^{-1}(\text{dom } K) \subset \text{dom } KA$. The inclusion can be strengthened to equality if K is closed and proper and range $A \cap \text{ri}(\text{dom } K) \neq \emptyset$.*

PROOF. Write $\text{dom } K = C \times D$. By (5.7), KA is concave-convex. If $u \in A_1^{-1}C$, then $K(A_1u,\cdot)$ is never $-\infty$ and hence $KA(u,\cdot) = K(A_1u,\cdot)A_2$ is never $-\infty$. This shows $A_1^{-1}C \subset \text{dom}_1 KA$. Similarly $A_2^{-1}D \subset \text{dom}_2 KA$. Now assume K is closed and proper and range $A \cap \text{ri}(\text{dom } K) \neq \emptyset$. If $u \notin A^{-1}C$, then (34.3) implies $K(A_1u,\cdot)A_2$ equals $-\infty$ everywhere on ri D and hence $K(A_1u,\cdot)A_2$ equals $-\infty$ everywhere on $A_2^{-1}(\text{ri } D)$. Since $A_2^{-1}(\text{ri } D) \neq \emptyset$ by hypothesis, this shows $\text{dom}_1 KA \subset A_1^{-1}C$. Similarly $\text{dom}_2 KA \subset A_2^{-1}D$.

THEOREM 2.2. *Let* K *be a closed proper concave-convex function on* $R^m \times R^n$, *and let* $A = A_1 \times A_2$ *be a linear transformation from* $R^p \times R^q$ *to* $R^m \times R^n$ *such that* range $A \cap \text{ri}(\text{dom } K) \neq \emptyset$. *Then the collection* $\{\tilde{K}A | \tilde{K} \in [K]\}$ *of saddle-functions is contained in an equivalence class* $[H]$ *of closed proper concave-convex functions on* $R^p \times R^q$ *having domain* $A^{-1}(\text{dom } K)$. *Moreover*,

$$\underline{H} = \underline{K}A, \quad \overline{H} = \overline{K}A,$$

$$\text{ri}(\text{dom } H) = A^{-1}\text{ri}(\text{dom } K),$$

$$\text{cl}(\text{dom } H) = A^{-1}\text{cl}(\text{dom } K).$$

PROOF. Lemma 2.1 implies $\underline{K}A$ and $\overline{K}A$ are proper concave-convex functions on $R^p \times R^q$ with domain $A^{-1}(\text{dom } K)$. From Theorem 0.1 (b) it is clear that a closed proper saddle-function is the least member of its equivalence class iff it is convex-closed. Now it follows routinely, using (6.7), (34.3) and (9.5), that $\underline{K}A$ satisfies the six conditions of (34.3) and moreover is convex-closed. Hence $\underline{K}A$ is closed and is the least member of its equivalence class. Similarly, $\overline{K}A$ is closed and is the greatest member of its equivalence class. According to (34.4), two closed proper saddle-functions are equivalent iff they have the same kernel. Suppose $(u,v) \in \text{ri}(A^{-1}(\text{dom } K))$. By (6.7) this means $A(u,v) \in \text{ri}(\text{dom } K)$. Since $\underline{K}$ and $\overline{K}$ are equivalent closed proper, $\underline{K}A(u,v) = \overline{K}A(u,v)$. This shows $\underline{K}A \sim \overline{K}A$ and hence $[\underline{K}A] = [\overline{K}A] = [H]$. If $K \in [K]$, then $\underline{K} \leq K \leq \overline{K}$ implies $\underline{K}A \leq KA \leq \overline{K}A$ and hence $KA \in [H]$ (Theorem 0.1(b)). The formulas for $\text{ri}(\text{dom } H)$ and $\text{cl}(\text{dom } H)$ are immediate from (6.7).

THEOREM 2.3. *Let* K *and* A *be as in Theorem 2.2. Then*

$$\partial(KA)(u,v) = A^*\partial K(A(u,v))$$

for each $(u,v) \in R^p \times R^q$, *and*

$$\text{ri}(A^{-1}(\text{dom } K)) \subset \text{dom } \partial(KA) \subset A^{-1}(\text{dom } K).$$

PROOF. The inclusions are immediate from (37.4) and Theorem 2.2. It follows that the identity holds trivially when $(u,v) \notin A^{-1}(\text{dom } K)$. Suppose $(u,v) \in A^{-1}$ dom K. By the definitions, $(u^*,v^*) \in \partial(KA)(u,v)$ iff $u^* \in \partial(K(\cdot,A_2v)A_1)(u)$ and $v^* \in \partial(K(A_1u,\cdot)A_2)(v)$. Now by (34.3), $A_1u \in \text{dom}_1K$ implies that $K(A_1u,\cdot)$ is a proper convex function with $\text{ri}(\text{dom } K(A_1u,\cdot)) = \text{ri}(\text{dom}_2K)$. Hence range $A_2 \cap \text{ri}(\text{dom}_2K) \neq \emptyset$ and (23.9) imply that $v^* \in \partial(K(A_1u,\cdot)A_2)(v)$ iff $v^* \in A_2^*\partial K(A_1u,\cdot)(A_2v)$, i.e. iff $v^* \in A_2^*\partial_2K(A(u,v))$. Similarly, $u^* \in \partial(K(\cdot,A_2v)A_1)(u)$ iff $u^* \in A_1^*\partial_1K(A(u,v))$. The identity follows.

THEOREM 2.4. *Let* K *and* A *be as in Theorem* 2.2. *Let* f (*resp.* g) *denote the convex* (*resp. concave*) *parent of* K, *and let* h (*resp.* k) *denote the convex* (*resp. concave*) *parent of* KA. *Then*

$$h(u,v^*) = (A_2^*f(A_1u,\cdot))(v^*)$$

and

$$k(u^*,v) = (A_1^*g(\cdot,A_2v))(u^*).$$

PROOF. Suppose $u \notin \text{dom}_1KA$. Then $h(u,\cdot)$ is constantly $+\infty$. Also, $A_1u \notin \text{dom}_1K$ implies $f(A_1u,\cdot)$ is constantly $+\infty$ and hence

$$(A_2^*f(A_1u,\cdot))(v^*) = \inf\{f(A_1u,y^*)\,|\,A_2^*y^* = v^*\} = +\infty$$

for every v^*. Now suppose $u \in \text{dom}_1KA$. By (34.3), $A_1u \in \text{dom}_1K$ implies $K(A_1u,\cdot)$ is a proper convex function with $\text{ri}(\text{dom } K(A_1u,\cdot)) = \text{ri}(\text{dom}_2K)$. Thus from (16.3), Theorem 0.1 (a) and range $A_2 \cap \text{ri}(\text{dom}_2K) \neq \emptyset$ it follows that $h(u,v^*) = (K(A_1u,\cdot)A_2)^*(v^*) = (A_2^*K(A_1u,\cdot)^*)(v^*) = (A_2^*f(A_1u,\cdot))(v^*)$ for every v^*. This proves the first identity. The second is proved similarly.

COROLLARY 2.4.1. *Let* K *and* A *be as in Theorem* 2.2. *If* K *is polyhedral, then* KA *is polyhedral.*

PROOF. Let I_p and I_n denote the identity transformations on R^p and R^n, respectively, and let f and h be as in Theorem 2.4. Then

$h = (I_p \times A_2^*)(f(A_1 \times I_n))$. Hence (19.3.1) implies that h is polyhedral if f is polyhedral. Since KA is closed by Theorem 2.2, this concludes the proof.

COROLLARY 2.4.2. Let K and A be as in Theorem 2.2. Then

$$\mathrm{dom}(KA)^* \subset A^*\mathrm{dom}\, K^*.$$

In particular, if f (resp. g) denotes the convex (resp. concave) parent of K, then

$$\mathrm{ri}(\mathrm{dom}_1(KA)^*) = A_1^* \bigcup\{\mathrm{ri}(\mathrm{dom}\, g(\cdot,y)) \mid y \in \mathrm{range}\, A_2 \cap \mathrm{ri}(\mathrm{dom}_2 K)\}$$

and

$$\mathrm{ri}(\mathrm{dom}_2(KA)^*) = A_2^* \bigcup\{\mathrm{ri}(\mathrm{dom}\, f(x,\cdot)) \mid x \in \mathrm{range}\, A_1 \cap \mathrm{ri}(\mathrm{dom}_1 K)\},$$

where these formulas also hold when "ri" is deleted throughout.

PROOF. By Lemma 0.3.

LEMMA 2.5. Let K and A be as in Theorem 2.2. Then

$$(\mathrm{rec}_1 KA)(u) = \inf\{(\mathrm{rec}\, K(\cdot,y))(A_1 u) \mid y \in \mathrm{range}\, A_2 \cap \mathrm{ri}(\mathrm{dom}_2 K)\}$$

and

$$(\mathrm{rec}_2 KA)(v) = \sup\{(\mathrm{rec}\, K(x,\cdot))(A_2 v) \mid x \in \mathrm{range}\, A_1 \cap \mathrm{ri}(\mathrm{dom}_1 K)\}.$$

PROOF. By definition and Theorem 2.2,

$$\begin{aligned}(\mathrm{rec}_2 KA)(v) &= \sup\{(\mathrm{rec}(KA)(u,\cdot))(v) \mid u \in A_1^{-1}\mathrm{ri}(\mathrm{dom}_1 K)\}.\\ &= \sup\{(\mathrm{rec}\, K(x,\cdot)A_2)(v) \mid x \in \mathrm{range}\, A_1 \cap \mathrm{ri}(\mathrm{dom}_1 K)\}.\end{aligned}$$

If $x \in \mathrm{ri}(\mathrm{dom}_1 K)$, then (34.3) and (9.5) imply $(\mathrm{rec}\, K(x,\cdot)A_2)(v) = (\mathrm{rec}\, K(x,\cdot))(A_2 v)$. This proves one formula, and the other is proved similarly.

THEOREM 2.6. Let K and A be as in Theorem 2.2. Then for $j =$ 1 and 2,

$$\mathrm{cl}(\mathrm{dom}_j(KA)^*) = \mathrm{cl}(A_j^*\mathrm{dom}_j K^*)$$

iff

$$\mathrm{rec}_j(KA) = (\mathrm{rec}_j K)A_j.$$

PROOF. By (37.2), $\mathrm{rec}_2(KA) = \delta^*(\cdot \mid \mathrm{dom}_2(KA)^*)$ and

$(\mathrm{rec}_2K)A_2 = \delta^*(A_2\cdot|\mathrm{dom}_2K^*)$. Now apply (16.3.1) and (13.1.1). The assertion for $j = 1$ is proved similarly.

Next we develop an operation analogous to that of taking the image Af of a convex function f under a linear transformation A. Suppose K is a concave-convex function on $R^m \times R^n$ and $A = A_1 \times A_2$ is a linear transformation from $R^m \times R^n$ to $R^p \times R^q$.

We seek a condition on K and A ensuring that all the functions on $R^p \times R^q$ either of the form

$$(u,v) \to \sup_{\{x|A_1x = u\}} \inf_{\{y|A_2y = v\}} \tilde{K}(x,y)$$

or of the form

$$(u,v) \to \inf_{\{y|A_2y = v\}} \sup_{\{x\ A_1x = u\}} \tilde{K}(x,y),$$

for $\tilde{K} \in [K]$, belong to a single equivalence class of concave-convex functions on $R^p \times R^q$. By analogy with the operation in the convex function case, this equivalence class (when it exists) will usually be denoted by $[AK]$. Theorem 2.8 gives a condition which guarantees that $[AK]$ exists and, moreover, that all of its members are closed and proper, and that its conjugate is $[K^*A^*]$.

LEMMA 2.7. <u>Let</u> K <u>be a closed proper concave-convex function on</u> $R^m \times R^n$, <u>and let</u> $A = A_1 \times A_2$ <u>be a linear transformation from</u> $R^m \times R^n$ <u>to</u> $R^p \times R^q$. <u>Assume</u> range $A^* \cap \mathrm{ri}(\mathrm{dom}\ K^*) \neq \emptyset$. <u>Then</u>

$$(\overline{K^*A^*})^*(u,v) = \sup\{\mathrm{cl}(A_2\underline{K}(x,\cdot))(v)\,|\,A_1x = u\},$$

<u>where the supremum can be taken over just those</u> x <u>in</u> dom_1K <u>such that</u> $A_1x = u$, <u>and</u>

$$(\underline{K^*A^*})^*(u,v) = \inf\{\mathrm{cl}(A_1\overline{K}(\cdot,y))(u)\,|\,A_2y = v\},$$

<u>where the infimum can be taken over just those</u> y <u>in</u> dom_2K <u>such that</u> $A_2y = v$.

PROOF. Only the first formula will be proved, as the second can be proved similarly. Let J denote the lower conjugate of $\overline{K}^*A^*$. The definitions yield

$$J(u,v) = \sup_{v^*}\{\langle v^*,v\rangle + \inf_{u^*}\{\langle u^*,u\rangle - (\overline{K}^*(\cdot,A_2^*v^*)A_1^*)(u^*)\}\}.$$

Since $\overline{K}^*$ is concave-closed, it follows from (34.3) and (6.3.1) that $\mathrm{ri}(\mathrm{dom}\,\overline{K}^*(\cdot,y^*))$ equals $\mathrm{ri}(\mathrm{dom}_1 K^*)$ when $y^* \in \mathrm{dom}_2 K^*$ and equals R^m when $y^* \notin \mathrm{dom}_2 K^*$. Hence (16.3) and the hypothesis range $A_1^* \cap \mathrm{ri}(\mathrm{dom}_1 K^*) \neq \emptyset$ imply $(\overline{K}^*(\cdot,A_2^*v^*)A_1^*)^*(u) = (A_1\overline{K}^*(\cdot,A_2^*v^*)^*)(u) = \sup\{k(x,A_2^*v^*)\,|\,A_1x = u\}$ for every v^*, where k denotes the concave relative of K^*. Thus,

$$\begin{aligned} J(u,v) &= \sup_{v^*}\{\langle v^*,v\rangle + \sup_{x\in A_1^{-1}u} k(x,A_2^*v^*)\} \\ &= \sup_{x\in A_1^{-1}u}\,\sup_{v^*}\{\langle v^*,v\rangle - (-k)(x,A_2^*v^*)\}. \end{aligned}$$

But Theorem 0.2 implies $-k$ is the convex parent f of K, and hence (16.3) implies $\sup_{v^*}\{\langle v^*,v\rangle - (-k)(x,A_2^*v^*)\} = (f(x,\cdot)A_2^*)^*(v) = \mathrm{cl}(A_2 f(x,\cdot)^*)(v) = \mathrm{cl}(A_2\underline{K}(x,\cdot))(v)$. This establishes the asserted formula for J. Finally, for each $x \notin \mathrm{dom}_1 K$, the fact that $\underline{K}$ is convex-closed implies $\underline{K}(x,\cdot)$ and hence $\mathrm{cl}(A_2\underline{K}(x,\cdot))$ is constantly $-\infty$.

THEOREM 2.8. *Let K and A be as in Lemma 2.7 and assume range $A^* \cap \mathrm{ri}(\mathrm{dom}\,K^*) \neq \emptyset$. Define functions J_1 and J_2 on $R^p \times R^q$ by*

$$J_1(u,v) = \sup_{\{x|A_1x = u\}}\ \inf_{\{y|A_2y = v\}} \underline{K}(x,y)$$

and

$$J_2(u,v) = \inf_{\{y|A_2y = v\}}\ \sup_{\{x|A_1x = u\}} \overline{K}(x,y).$$

*Then there exists an equivalence class [AK] which contains every concave-convex function J on $R^p \times R^q$ satisfying $J_1 \le J \le J_2$. Moreover, [AK] is closed and proper and its conjugate is [K*A*]. If [K] is polyhedral, then [AK] is polyhedral.*

PROOF. Theorem 2.2 implies that $\underline{K}^*A^*$ and $\overline{K}^*A^*$ belong to a closed proper equivalence class $[K^*A^*]$. Let $[AK]$ denote the conjugate equivalence class. From Lemma 2.7 and the fact that $\mathrm{cl}\, f \leq f$ for any convex function f and $\mathrm{cl}\, g \geq g$ for any concave function g, it follows that

$$(\underline{K^*A^*})^* \leq J_1 \quad \text{and} \quad J_2 \leq (\overline{K^*A^*})^*$$

Hence Theorem 0.1(b) implies that each concave-convex function J on $R^p \times R^q$ satisfying $J_1 \leq J \leq J_2$ belongs to $[AK]$. The polyhedral assertion follows from Corollary 2.4.1 and the fact that K^* is polyhedral whenever K is.

The following lemma dualizes the hypothesis used throughout this section.

LEMMA 2.9. _Let_ K _and_ A _be as in Lemma 2.7. Then for_ $j = 1$ _and_ 2 _the following conditions are equivalent_:

(a_j) range $A_j^* \cap \mathrm{ri}(\mathrm{dom}_j K^*) \neq \emptyset$;

(b_j) $A_j^{-1}(0) \cap (\mathrm{rec\ cone}_j K)$ _is a subspace_;

(c_j) $A_j^{-1}(0) \cap (\mathrm{rec\ cone}_j K) \subset -(\mathrm{rec\ cone}_j K)$.

PROOF. Apply Lemma 0.5 with $L_j = \mathrm{range}\ A_j^*$.

We conclude this section with two examples showing that Theorems 2.2 and 2.8 can fail if their hypotheses are weakened. These examples are presented in the notational scheme of Theorems 1.1 and 2.2.

EXAMPLE 2.10. Take $m = n = p = q = 1$, and let A_1 and A_2 each be the zero transformation on R. Let K be a member of the equivalence class of closed proper concave-convex functions on $R \times R$ having as kernel the function

$$(u,v) \to u^v,\ \forall (u,v) \in (0,1) \times (0,1).$$

(This equivalence class is discussed in [44, p. 360].) Then $\mathrm{dom}\, K = [0,1] \times [0,1]$, $\underline{K}(0,0) = 0$, $\overline{K}(0,0) = 1$, and $\underline{K}(u,v) = u^v = \overline{K}(u,v)$ whenever

$(u,v) \in \text{dom } K \setminus \{(0,0)\}$. Moreover, for each $\alpha \in [0,1]$ the function K_α belongs to $[K]$, where $K_\alpha(0,0) = \alpha$ and $K_\alpha(u,v) = \underline{K}(u,v)$ whenever $(u,v) \neq (0,0)$. Observe that, for $j = 1$ and 2, range $A_j \cap \text{dom}_j K \neq \emptyset$ while range $A_j \cap \text{ri}(\text{dom}_j K) = \emptyset$. Also, for any $\tilde{K} \in [K]$, the function $\tilde{K}A$ is constantly equal to $\tilde{K}(0,0)$. Since $0 \leq \tilde{K}(0,0) \leq 1$, this implies that $\tilde{K}A$ is closed and proper. However, it also implies that, for any two elements K_1 and K_2 of $[K]$, K_1A is equivalent to K_2A iff $K_1(0,0) = K_2(0,0)$. Thus, as $\tilde{K}$ ranges over $[K]$ the functions $\tilde{K}A$ determine $2^{\aleph_0}$ distinct equivalence classes of closed proper saddle-functions (cf. Theorem 2.2). Now let J_1 and J_2 be as in Theorem 1.1. Since A_j^* is the zero transformation on R,

$$J_1(u^*,v^*) = \begin{cases} \sup\limits_R \inf\limits_R \underline{K}^* & \text{if } u^* = 0 \text{ and } v^* = 0 \\ +\infty & \text{if } u^* = 0 \text{ and } v^* \neq 0 \\ -\infty & \text{if } u^* \neq 0 \end{cases}$$

and

$$J_2(u^*,v^*) = \begin{cases} \inf\limits_R \sup\limits_R \bar{K}^* & \text{if } u^* = 0 \text{ and } v^* = 0 \\ -\infty & \text{if } u^* \neq 0 \text{ and } v^* = 0 \\ +\infty & \text{if } \qquad\qquad v^* \neq 0. \end{cases}$$

But $\sup\limits_R \inf\limits_R \underline{K}^* = -(\underline{K}^*)^*(0,0) = -\bar{K}(0,0) = -1$, and similarly $\inf\limits_R \sup\limits_R \bar{K}^* = 0$. Hence J_1 and J_2 are closed and proper but not equivalent (cf. Theorem 2.8).

EXAMPLE 2.11. Let K and A_1 be as in Example 2.10, but now let A_2 be the identity transformation on R. Observe that range $A_2 \cap \text{ri}(\text{dom}_2 K) \neq \emptyset$ and range $A_1 \cap \text{dom}_1 K \neq \emptyset$ but range $A_1 \cap \text{ri}(\text{dom}_1 K) = \emptyset$. For each $\tilde{K} \in [K]$,

$$\tilde{K}A(u,v) = \tilde{K}(0,v) = \begin{cases} 0 & \text{if } v \in (0,1] \\ \tilde{K}(0,0) & \text{if } v = 0 \\ +\infty & \text{if } v \notin [0,1], \end{cases}$$

where $0 \leq \tilde{K}(0,0) \leq 1$. This implies that $\tilde{K}A$ is proper with domain

$R \times [0,1]$ and that $\tilde{K}A$ is closed iff $\tilde{K}(0,0) = 0$. It also implies that, for any two elements K_1 and K_2 of $[K]$, K_1A is equivalent to K_2A iff $K_1(0,0) = K_2(0,0)$. Recalling the functions $K\alpha$ for $\alpha \in [0,1]$, we conclude that as $\tilde{K}$ ranges over $[K]$ the functions $\tilde{K}A$ determine $2^{\aleph_0}$ distinct equivalence classes of proper saddle-functions, where only the class containing $\underline{K}A$ is closed (cf. Theorem 2.2).

§3: Sharper Results

In this section Theorem 1.1, which has already been sharpened in §2, is sharpened still further. The principal results are Theorems 3.4 and 3.5. Among the conclusions are facts concerning the attainment of the extrema appearing in the definitions of J_1 and J_2. Lemma 3.6 states simple conditions sufficient for the more general hypotheses of Theorems 3.4 and 3.5 to hold.

Throughout §3 we adopt the notational setting of Theorem 2.8. That is, K is a closed proper concave-convex function on $R^m \times R^n$, $A = A_1 \times A_2$ is a linear transformation from $R^m \times R^n$ to $R^p \times R^q$, and J_1 and J_2 are functions defined on $R^p \times R^q$ by

$$J_1(u,v) = \sup_{\{x|A_1x = u\}} \inf_{\{y|A_2y = v\}} \underline{K}(x,y)$$

and

$$J_2(u,v) = \inf_{\{y|A_2y = v\}} \sup_{\{x|A_1x = u\}} \overline{K}(x,y).$$

LEMMA 3.1. *Let f be a proper convex function on R^n, let D be a convex set such that $D \subset \text{dom } f \subset \text{cl } D$, and let E be a convex set such that $E \cap \text{ri } D \neq \emptyset$. Then*

$$\inf_{E \cap D} f = \inf_E f.$$

PROOF. By (6.3.1), $D \subset \text{dom } f \subset \text{cl } D$ implies $\text{ri } D = \text{ri}(\text{dom } f)$. Hence $S = E \cap \text{ri}(\text{dom } f) \subset E \cap D \subset E$ implies trivially that

$$\inf_S f \geq \inf_{E \cap D} f \geq \inf_E f.$$

Let $y \in E$ be given. If $y \notin \text{dom } f$, then $f(y) = +\infty \geq \inf_S f$. Suppose $y \in \text{dom } f$. Since $E \cap \text{ri } D \neq \emptyset$, we can pick an $x \in S$. Then (6.1) implies that $y_\lambda = (1 - \lambda)x + \lambda y \in S$ for each $0 \leq \lambda < 1$. Hence (7.5) implies that $f(y) \geq \text{cl } f(y) = \lim_{\lambda \uparrow 1} f(y_\lambda) \geq \inf\{f(y_\lambda) | 0 \leq \lambda < 1\} \geq \inf_S f$. This shows that $f(y) \geq \inf_S f$ for every $y \in E$. Thus $\inf_E f \geq \inf_S f$, and the proof is

complete.

THEOREM 3.2. _Let_ $(u,v) \in A\,\mathrm{ri}(\mathrm{dom}\,K)$ _and assume that_

$$A_2^{-1}(0) \cap \bigcap \{\text{rec cone } \underline{K}(x,\cdot) \mid x \in \mathrm{ri}(\mathrm{dom}_1 K),\ A_1 x = u\}$$

and

$$A_1^{-1}(0) \cap \bigcap \{\text{rec cone } \overline{K}(\cdot,y) \mid y \in \mathrm{ri}(\mathrm{dom}_2 K),\ A_2 y = v\}$$

are subspaces. Then there exists a nonempty closed convex product set $X \times Y$ _in_ $\mathrm{dom}\,\partial K \cap A^{-1}(\{(u,v)\})$ _such that_ $(x,y) \in X \times Y$ _iff_ (x,y) _is a saddle-point of_ K _with respect to_ $A^{-1}(\{(u,v)\})$ _for each_ $\tilde{K} \in [K]$. _If the two sets in the hypothesis are actually nullspaces, then_ $X \times Y$ _is bounded._

PROOF. Define a concave-convex function L on $R^m \times R^n$ by

$$L(x,y) = \begin{cases} 0 & \text{if } A_1 x = u \text{ and } A_2 y = v \\ +\infty & \text{if } A_1 x = u \text{ and } A_2 y \neq v \\ -\infty & \text{if } A_1 x \neq u \end{cases} .$$

Clearly, L is closed and has domain $A^{-1}\{(u,v)\}$. Since $(u,v) \in A\,\mathrm{ri}(\mathrm{dom}\,K)$, $\mathrm{ri}(\mathrm{dom}\,K) \cap \mathrm{ri}(\mathrm{dom}\,L) \neq \emptyset$. Therefore Theorem 4.2 (which doesn't depend on the results of §3) implies that the equivalence class $[K] + [L]$ is defined and has domain $S \times T$, where $S = A_1^{-1}\{u\} \cap \mathrm{dom}_1 K$ and $T = A_2^{-1}\{v\} \cap \mathrm{dom}_2 K$. Moreover, Theorem 4.2 also implies that for any $\tilde{K} \in [K]$, $[K] + [L]$ contains the closed proper saddle-function M given by

$$M(x,y) = \begin{cases} \tilde{K}(x,y) & \text{if } x \in S \text{ and } y \in T \\ +\infty & \text{if } x \in S \text{ and } y \notin T \\ -\infty & \text{if } x \notin S \end{cases} .$$

Suppose $x \in \mathrm{ri}\,S = A_1^{-1}\{u\} \cap \mathrm{ri}(\mathrm{dom}_1 K)$ (use [6.5]). Then (34.3) implies $\tilde{K}(x,\cdot) = \underline{K}(x,\cdot)$ is a closed proper convex function with effective domain $\mathrm{dom}_2 K$. Hence (9.3) and the definition of M imply

$$\text{rec } M(x,\cdot) = \text{rec } \underline{K}(x,\cdot) + \text{rec } \delta(\cdot \mid A_2^{-1}(v)).$$

But $\text{rec } \delta(\cdot \mid A_2^{-1}(v)) = \delta(\cdot \mid A_2^{-1}(0))$. Therefore

$$\text{rec } M(x,\cdot) = A_2^{-1}\{0\} \cap \text{rec cone } \underline{K}(x,\cdot).$$

Since $M(x,\cdot) = \underline{M}(x,\cdot)$ whenever $x \in \text{ri } S$ (Theorem 0.1(b)), this implies that

$$\text{rec cone}_2 M = A_2^{-1}\{0\} \cap \bigcap \{\text{rec cone } \underline{K}(x,\cdot) \mid x \in \text{ri } S\}.$$

By hypothesis this is a subspace. Similarly, $\text{rec cone}_1 M$ is a subspace. It follows from Lemma 0.4 that $(0,0) \in \text{ri}(\text{dom } M^*)$, and hence (37.5.3) implies that $\partial M^*(0,0)$ is a nonempty closed convex product set $X \times Y$. By (37.5), $\partial M^*(0,0) \subset \text{dom } \partial M$. But Theorem 4.9 (which doesn't depend on the results of §3) implies that $\text{dom } \partial M = \text{dom } \partial K \cap \text{dom } \partial L$, and (37.4) implies $\text{dom } \partial L \subset \text{dom } L$. Therefore $X \times Y \subset \text{dom } \partial K \cap A^{-1}\{(u,v)\}$. Now $(x,y) \in X \times Y$ iff (x,y) is a saddle-point of M, which (by (36.3)) occurs iff (x,y) is a saddle-point of $\tilde{K}$ with respect to $S \times T$. Using $(x,y) \in \text{dom } K$ together with (34.3) and Lemma 3.1, it follows that this occurs iff (x,y) is a saddle-point of $\tilde{K}$ with respect to $A^{-1}\{(u,v)\}$. Since any member of $[K]$ could be taken in the definition of M, (36.4) implies that $(x,y) \in X \times Y$ iff (x,y) is a saddle-point of $\tilde{K}$ with respect to $A^{-1}\{(u,v)\}$ for each $\tilde{K} \in [K]$. Finally, suppose the sets in the hypothesis are actually nullspaces. Then $\text{rec cone}_j M = \{0\}$ for $j = 1$ and 2, so that Lemma 0.4 implies $(0,0) \in \text{int}(\text{dom } M^*)$. From this, (34.3) and (23.4) it follows that the sets $\partial M^*(\cdot,0)(0) = X$ and $\partial M^*(0,\cdot)(0) = Y$ are bounded.

LEMMA 3.3. *For $x \in \text{dom}_1 K$ the following three conditions are equivalent, and they imply $A_1 x \in \text{dom}_1 J_1$:*

(a$_1$) $\text{range } A_2^* \cap \text{ri}(\text{dom } K(x,\cdot)^*)$;

(a$_2$) $A_2^{-1}\{0\} \cap \text{rec cone } \underline{K}(x,\cdot)$ *is a subspace*;

(a$_3$) $A_2^{-1}\{0\} \cap \text{rec cone } \underline{K}(x,\cdot) \subset -\text{rec cone } \underline{K}(x,\cdot)$.

Similarly, for $y \in \text{dom}_2 K$ the following three conditions are equivalent and they imply $A_2 y \in \text{dom}_2 J_2$:

(b_1) range $A_1^* \cap$ ri(dom $K(\cdot,y)^*$);

(b_2) $A_1^{-1}(0) \cap$ rec cone $\overline{K}(\cdot,y)$ is a subspace;

(b_3) $A_1^{-1}(0) \cap$ rec cone $\overline{K}(\cdot,y) \subset$ -rec cone $\overline{K}(\cdot,y)$.

PROOF. Only the first assertion is proved, as the second can be proved similarly. Since rec cone $\underline{K}(x,\cdot)$ is a convex cone, clearly (a_2) holds iff (a_3) holds. By Theorem 0.1, $K(x,\cdot)^*$ is proper convex and its conjugate is $\underline{K}(x,\cdot)$. Hence (16.2.1) implies that (a_1) fails iff (a_3) fails. Thus, the three conditions (a_1) - (a_3) are equivalent. Suppose now that x satisfies (a_3). Since $\underline{K}(x,\cdot)$ is closed proper convex, (9.2) implies that $A_2\underline{K}(x,\cdot)$ is too. Hence $A_2\underline{K}(x,\cdot)$ is never $-\infty$. But $A_2\underline{K}(x,\cdot) \le J_1(A_1x,\cdot)$. Therefore $A_1x \in \mathrm{dom}_1 J_1$.

THEOREM 3.4. Assume that each $x \in$ ri($\mathrm{dom}_1 K$) (resp. $y \in$ ri($\mathrm{dom}_2 K$)) satisfies one of the equivalent conditions (a_i) (resp. (b_i)) of Lemma 3.3. Then the conclusions of Theorem 2.8 hold, and

$$\mathrm{ri}(A\ \mathrm{dom}\ K) \subset \mathrm{dom}\ AK \subset A\ \mathrm{dom}\ K.$$

Furthermore, for each $(u,v) \in$ ri(A dom K) there exists a nonempty closed convex product set $X \times Y$ in dom $\partial K \cap A^{-1}\{(u,v)\}$ such that (1) $(x,y) \in X \times Y$ iff (x,y) is a saddle-point of $\tilde{K}$ with respect to $A^{-1}\{(u,v)\}$ for each $\tilde{K} \in [K]$, and (2) $(x,y) \in X \times Y$ implies $\tilde{J}(u,v) = \tilde{K}(x,y)$ for every $\tilde{J} \in [AK]$ and $\tilde{K} \in [K]$.

PROOF. The hypothesis implies that $A_j^{-1}(0) \cap$ rec $\mathrm{cone}_j K$ is a subspace for $j = 1$ and 2. Hence by Lemma 2.9 the conclusions of Theorem 2.8 hold, and in particular J_1 and J_2 belong to [AK], where $[(AK)^*] = [K^*A^*]$. Thus dom AK $= \mathrm{dom}_1 J_1 \times \mathrm{dom}_2 J_2$. Therefore the hypothesis and Lemma 3.3 imply that ri(A dom K) $\subset$ dom AK. On the other hand, Lemma 2.9 and Corollary 2.4.2 imply that dom AK $\subset$ A dom K. Let $(u,v) \in$ ri(A dom K). By Theorem 3.2 there exists a nonempty closed convex product set $X \times Y$ in

$\text{dom}\,\partial K \cap A^{-1}(\{(u,v)\})$ such that (1) holds. Suppose $(x,y) \in X \times Y$. Since (1) implies (x,y) is a saddle-point of $\underline{K}$ with respect to $A^{-1}(\{(u,v)\})$, certainly $J_1(u,v) = \underline{K}(x,y)$. Since $\text{ri}(A\,\text{dom}\,K) = \text{ri}(\text{dom}\,AK)$ by (6.3.1), Theorem 0.1(b) implies that $J_1(u,v) = \tilde{J}(u,v)$ for each $\tilde{J} \in [AK]$. Also, $(x,y) \in \text{dom}\,\partial K$ and (37.4.1) imply that $\underline{K}(x,y) = \tilde{K}(x,y)$ for each $\tilde{K} \in [K]$. This establishes (2).

THEOREM 3.5. <u>Assume that each</u> $x \in \text{dom}_1 K$ (<u>resp.</u> $y \in \text{dom}_2 K$) <u>satisfies one of the equivalent conditions</u> (a_1) (<u>resp.</u> (b_1)) <u>of Lemma 3.3. Then</u> dom AK <u>actually equals</u> A dom K. <u>Moreover, writing</u> $\text{cl}_2(AK) = \underline{J}$ <u>and</u> $\text{cl}_1(AK) = \bar{J}$,

$$J_1 = \underline{J} \ \text{and} \ J_2 = \bar{J} \ \text{on range } A.$$

<u>In particular,</u> $J_1(u,v) = \underline{J}(u,v)$ <u>except when</u> $u \in \text{range}\,A_1 \setminus A_1\text{dom}_1 K$ <u>and</u> $v \notin \text{range}\,A_2$, <u>and</u> $J_2(u,v) = \bar{J}(u,v)$ <u>except when</u> $u \notin \text{range}\,A_1$ <u>and</u> $v \in \text{range}\,A_2 \setminus A_2\text{dom}_2 K$.

PROOF. By Lemma 3.3, $A\,\text{dom}\,K \subset \text{dom}_1 J_1 \times \text{dom}_2 J_2$. From this it follows as in the proof of Theorem 3.4 that $A\,\text{dom}\,K = \text{dom}\,AK$. We only prove the assertion about J_1, as the other is similar. From the proof of Lemma 3.3, $A_2\underline{K}(x,\cdot)$ is closed for each $x \in \text{dom}_1 K$. Hence Lemma 2.7 implies

$$\underline{J}(u,v) = \sup\{A_2\underline{K}(x,\cdot)(v) \mid x \in \text{dom}_1 K,\ A_1 x = u\}. \tag{1}$$

If $x \notin \text{dom}_1 K$, then $\underline{K}(x,\cdot)$ is constantly $-\infty$, so that

$$A_2\underline{K}(x,\cdot)(v) = \begin{cases} -\infty & \text{if } v \in \text{range}\,A_2 \\ +\infty & \text{if } v \notin \text{range}\,A_2. \end{cases} \tag{2}$$

Since $J_1(u,v) = \sup\{A_2\underline{K}(x,\cdot)(v) \mid A_1 x = u\}$ by definition, (1) implies that $J_1(u,v)$ equals

$$\max\{\underline{J}(u,v),\ \sup\{A_2\underline{K}(x,\cdot)(v) \mid A_1 x = u,\ x \notin \text{dom}_1 K\}\},$$

which by (2) equals $\underline{J}(u,v)$ whenever $v \in \text{range}\,A_2$. Henceforth assume $v \notin \text{range}\,A_2$. Suppose $u \in A_1\text{dom}_1 K$. Pick an $x \in \text{dom}_1 K$ such that $A_1 x = u$. Since $A_2^{-1}(v) = \emptyset$, $+\infty = A_2\underline{K}(x,\cdot)(v) \le \underline{J}(u,v) \le J_1(u,v)$. Hence

$\underline{J}(u,v) = J_1(u,v) = +\infty$ whenever $u \in A_1\mathrm{dom}_1K$. Observe also that the convention $\sup \emptyset = -\infty$ implies $\underline{J}(u,v) \leq J_1(u,v) = -\infty$ whenever $u \notin \mathrm{range}\ A_1$. In the only remaining case, i.e. when $u \in \mathrm{range}\ A_1 \setminus A_1\mathrm{dom}_1K$, (1) implies $\underline{J}(u,v) = \sup \emptyset = -\infty$ while $J_1(u,v) = \sup\{\inf \emptyset | A_1x = u\} = +\infty$.

While the hypotheses of Theorems 3.4 and 3.5 are general, they may appear somewhat cumbersome to check. The next lemma gives simpler, "global" conditions on K and A which imply the hypotheses of both Theorems 3.4 and 3.5.

For a nonempty convex set C in R^n, define the <u>recession cone</u> of C to be the set

$$0^+ C = \{y | x + \lambda y \in C, \forall x \in C \ \forall \lambda \geq 0\}.$$

LEMMA 3.6. *The three following conditions are equivalent, and they imply that conditions* (a_1) - (a_3) *of Lemma 3.3 hold for each* $x \in \mathrm{dom}_1K$:

(c_1) $A_2^{-1}\{0\} \cap 0^+ \mathrm{cl}(\mathrm{dom}_2K) = \{0\}$;

(c_2) $A_2^{-1}\{v\} \cap \mathrm{dom}_2K$ *is bounded for each* $v \in R^q$;

(c_3) $A_2^{-1}\{v\} \cap \mathrm{ri}(\mathrm{dom}_2K)$ *is nonempty and bounded for some* $v \in R^q$.

Similarly, the three following conditions are equivalent, and they imply that conditions (b_1) - (b_3) *of Lemma 3.3 hold for each* $y \in \mathrm{dom}_2K$:

(d_1) $A_1^{-1}\{0\} \cap 0^+ \mathrm{cl}(\mathrm{dom}_1K) = \{0\}$;

(d_2) $A_1^{-1}\{u\} \cap \mathrm{dom}_1K$ *is bounded for each* $u \in R^p$;

(d_3) $A_1^{-1}\{u\} \cap \mathrm{ri}(\mathrm{dom}_1K)$ *is nonempty and bounded for some* $u \in R^p$.

PROOF. Only the first assertion is proved, since the second is similar. For each $v \in A_2\mathrm{dom}_2K$, (8.3.3) and (8.4) imply that

$$A_2^{-1}\{v\} \cap \mathrm{cl}(\mathrm{dom}_2K) \text{ is bounded iff } A_2^{-1}\{0\} \cap 0^+ \mathrm{cl}(\mathrm{dom}_2K) = \{0\}. \qquad (1)$$

It follows from this that (c_1) implies (c_2). By picking any $v \in A_2\mathrm{ri}(\mathrm{dom}_2K)$

it follows that (c_2) implies (c_3). Now assume (c_3). Then (6.3.1) and (6.5.1) imply that

$$A_2^{-1}\{v\} \cap \mathrm{cl}(\mathrm{dom}_2 K) = A_2^{-1}\{v\} \cap \mathrm{cl}(\mathrm{ri}(\mathrm{dom}_2 K)) = \mathrm{cl}(A_2^{-1}\{v\} \cap \mathrm{ri}(\mathrm{dom}_2 K)).$$

That this set is bounded follows from the fact that $A_2^{-1}\{v\} \cap \mathrm{ri}(\mathrm{dom}_2 K)$ is bounded. Hence (c_1) follows by (1). Finally, let $x \in \mathrm{dom}_1 K$ be given. Write $f = \underline{K}(x,\cdot)$ and $C = \mathrm{dom}\, f$. Then by (34.3) and (6.3.1), f is a proper convex function with $\mathrm{cl}\, C = \mathrm{cl}(\mathrm{dom}_2 K)$. But by (8.5) and (8.1) it follows easily that $\mathrm{dom}(\mathrm{rec}\, f) \subset 0^+ C \subset 0^+(\mathrm{cl}\, C)$. Hence $\mathrm{rec\ cone}\, \underline{K}(x,\cdot) \subset 0^+(\mathrm{cl}\, \mathrm{dom}_2 K)$, and therefore (c_1) implies that x satisfies (a_2) of Lemma 3.3.

Finally, observe that if conditions (c_1) and (d_1) of Lemma 3.6 are met, then the sets X and Y given by Theorem 3.4 for each $(u,v) \in A\, \mathrm{ri}(\mathrm{dom}\, K)$ are actually bounded and hence compact. This is because the two sets in the hypothesis of Theorem 3.2 are then nullspaces.

§4: Addition and Minimax Convolution

This section begins with the development of the addition operation on equivalence classes of saddle-functions. Next, some results concerning separable saddle-functions are presented. These are used together with the results of §§2 and 3 to define another operation on equivalence classes of saddle-functions. This operation, called minimax convolution, is dual to addition. The theorems proved concerning these dual operations closely parallel existing theorems about the dual operations of addition and infimal convolution on convex functions.

There are two technical difficulties involved in defining the operation of addition. The first stems from the fact that we are working with extended-real-valued functions; we must deal somehow with the expression $\infty - \infty$. The second and more fundamental difficulty is that, from the point of view of minimax theory, we want to define addition of whole equivalence classes and not just individual functions. The following definition is designed to handle both these difficulties.

For $i = 1,\ldots,s$ let K_i be a concave-convex function on $R^m \times R^n$ with domain $C_i \times D_i$. We say that $[K_1] +\ldots+ [K_s]$ <u>is defined</u> iff the sets $C = C_1 \cap \ldots \cap C_s$ and $D = D_1 \cap \ldots \cap D_s$ are nonempty and

$$\sum \tilde{K}_i(x,y) = \sum K_i(x,y), \quad \forall (x,y) \in \operatorname{ri} C \times \operatorname{ri} D$$

whenever $\tilde{K}_1 \in [K_1], \ldots, \tilde{K}_s \in [K_s]$. In this event $[K_1] +\ldots+ [K_s]$, which will usually be written as $[K_1 +\ldots+ K_s]$, or $[\sum K_i]$, is defined to be the unique equivalence class of closed proper concave-convex functions on $R^m \times R^n$ having as kernel the function on $\operatorname{ri} C \times \operatorname{ri} D$ given by $(x,y) \to \sum K_i(x,y)$. Such a unique equivalence class exists by (34.5.1). The operation which sends $[K_1], \ldots, [K_s]$ into $[K_1 +\ldots+ K_s]$ is, quite naturally, called <u>addition</u>.

LEMMA 4.1. <u>For</u> $i = 1,\ldots,s$ <u>let</u> K_i <u>be a closed proper concave-convex</u>

function on $R^m \times R^n$ with domain $C_i \times D_i$. Then $[K_1] +\ldots+ [K_s]$ is defined if either $C_1 \cap \ldots \cap C_s \neq \emptyset$ and $\text{ri } D_1 \cap \ldots \cap \text{ri } D_s \neq \emptyset$ or $\text{ri } C_1 \cap \ldots \cap \text{ri } C_s \neq \emptyset$ and $D_1 \cap \ldots \cap D_s \neq \emptyset$.

PROOF. This follows easily from (6.5) and Theorem 0.1(b).

It is actually not hard to establish a weaker condition sufficient for $[K_1] +\ldots+ [K_s]$ to be defined. Loosely speaking, the condition is just that the K_i be closed and that (possibly after renumbering the K_i's) there exist an integer r, $0 \leq r < s$, such that

$$\text{ri } C_1 \cap \ldots \cap \text{ri } C_r \cap C_{r+1} \cap \ldots \cap C_s \neq \emptyset$$

and

$$D_1 \cap \ldots \cap D_r \cap \text{ri } D_{r+1} \cap \ldots \cap \text{ri } D_s \neq \emptyset.$$

(The conditions in Lemma 4.1 correspond to the values $r = 0$ and $r = s - 1$.) Instead of appealing to (6.5), the proof uses the generalization of (6.5) described in the Appendix.

THEOREM 4.2. Let $K_1,\ldots,K_s$ be closed proper concave-convex functions on $R^m \times R^n$ such that $\text{ri}(\text{dom } K_1) \cap \ldots \cap \text{ri}(\text{dom } K_s) \neq \emptyset$. Then $[K_1] +\ldots+ [K_s]$ is defined, has domain $\text{dom } K_1 \cap \ldots \cap \text{dom } K_s$, and contains the closed proper saddle-function K given by

$$K(x,y) = \begin{cases} \sum K_i(x,y) & \text{if } x \in C \text{ and } y \in D \\ +\infty & \text{if } x \in C \text{ and } y \notin D \\ -\infty & \text{if } x \notin C \end{cases}.$$

PROOF. Lemma 4.1 implies $[K_1] +\ldots+ [K_s]$ is defined. Hence it is the unique equivalence class of closed proper concave-convex functions on $R^m \times R^n$ having the same kernel as K. Therefore by (34.4) the proof will be complete once we show K is closed. This we do by checking that K satisfies the six conditions of (34.3). This follows routinely by applying (34.3) to the K_i's with the aid of (6.5).

In order to apply the results of §§2 and 3 to an equivalence class $[K_1 + \ldots + K_s]$ and its conjugate, we need to define and establish some properties of "separable" saddle-functions. For $i = 1,\ldots,s$ let K_i be a proper concave-convex function on $R^{m_i} \times R^{n_i}$ with domain $C_i \times D_i$. Write $m = \sum m_i$, $n = \sum n_i$ and define a function $(K_1,\ldots K_s)$ on $R^m \times R^n$ by

$$(K_1,\ldots,K_s) = \begin{cases} \sum K_i(x_i,y_i) & \text{if } x \in C \text{ and } y \in D \\ +\infty & \text{if } x \in C \text{ and } y \notin D \\ -\infty & \text{if } x \notin C \end{cases}$$

where $x = (x_1,\ldots,x_s)$, $y = (y_1,\ldots,y_s)$, $C = C_1 \times\ldots\times C_s$, $D = D_1 \times\ldots\times D_s$. With the aid of (34.3) and the following Lemma 4.3, it can easily be verified that the function $(K_1,\ldots,K_s)$ is concave-concave with domain $C \times D$. Such a saddle-function is called <u>separable</u>. Lemma 4.3 reviews some useful properties of separable convex functions, and in Theorem 4.4 similar properties are established for separable saddle-functions.

LEMMA 4.3. *For $i = 1,\ldots,s$ let f_i be a proper convex function on R^{n_i} with effective domain C_i. Define $C = C_1 \times\ldots\times C_s$ and $f(x_1,\ldots,x_s) = f_1(x_1) +\ldots+ f_s(x_s)$. Then the following statements hold:*

(a) f *is proper convex with effective domain* C;

(b) $(\text{cl } f)(x_1,\ldots,x_s) = (\text{cl } f_1)(x_1) +\ldots+ (\text{cl } f_s)(x_s)$;

(c) f *is polyhedral if each* f_i *is*;

(d) $f^*(x_1^*,\ldots,x_s^*) = f_1^*(x_1^*) +\ldots+ f_s^*(x_s^*)$;

(e) $\partial f(x_1,\ldots,x_s) = \partial f_1(x_1) \times\ldots\times \partial f_s(x_s)$;

(f) $(\text{rec } f)(y_1,\ldots,y_s) = (\text{rec } f_1)(y_1) +\ldots+ (\text{rec } f_s)(y_s)$.

PROOF. Assertions (a) and (d) are trivial. Assertion (f) follows immediately from (a) and (8.5). To see (b), let $x = (x_1,\ldots,x_s) \in \text{cl } C = \text{cl } C_1 \times\ldots\times \text{cl } C_s$ be given and fix any $x_0^0 = (x_1,\ldots,x_s^0) \in \text{ri } C = \text{ri } C_1 \times\ldots\times \text{ri } C_s$. Define $x_\lambda = (x_1^\lambda,\ldots,x_s^\lambda)$ by $x_\lambda = (1 - \lambda)x_0 + \lambda x$ for

$0 \leq \lambda < 1$. Then (a) and (7.5) imply that

$$(\mathrm{cl}\ f)(x) = \lim_{\lambda \uparrow 1} f(x_\lambda) = \sum \lim_{\lambda \uparrow 1} f_i(x_i^\lambda) = \sum (\mathrm{cl}\ f_i)(x_i).$$

On the other hand, if $x \notin \mathrm{cl}\ C$ then $x_j \notin \mathrm{cl}\ C_j$ for some $1 \leq j \leq s$ and hence (a) and (7.4) imply that $(\mathrm{cl}\ f)(x) = +\infty = (\mathrm{cl}\ f_j)(x_j) \leq \sum (\mathrm{cl}\ f_i)(x_i)$. This proves (b). To see (c), define $h_i(x_1,\ldots,x_s) = f_i(x_i)$ for each i. Then

$$\mathrm{epi}\ h_i = \{(x_1,\ldots,x_s,\mu) \mid (x_i,\mu) \in \mathrm{epi}\ f_i\}$$

and $\mathrm{epi}\ f_i$ polyhedral imply that $\mathrm{epi}\ h_i$ is polyhedral for each i. Hence (19.4) implies that $f = h_1 + \ldots + h_s$ is polyhedral. Finally, we prove (e). Suppose first that $x = (x_1,\ldots,x_s) \notin C$. Then (a) and (23.4) imply that $\partial f(x) = \emptyset$ and also $\partial f_j(x_j) = \emptyset$ for some $1 \leq j \leq s$. Thus $\partial f_1(x_1) + \ldots + \partial f_s(x_s) = \emptyset$. Now suppose that $x \in C$. Using (6.1) and (7.5), one can easily verify that, for a convex function h on R^n and a subset C of R^n containing $\mathrm{ri}(\mathrm{dom}\ h)$, $x^* \in \partial h(x)$ iff

$$h(y) \geq h(x) + \langle x^*, y - x\rangle,\ \forall y \in C.$$

Applied to the situation at hand, this implies that $x^* = (x_1^*,\ldots x_s^*) \in \partial f(x)$ iff

$$\sum f_i(y_i) \geq \sum (f_i(x_i) + \langle x_i^*, y_i - x_i\rangle) \tag{1}$$

for every $(y_1,\ldots,y_s) \in C$. Let j be any fixed index. By letting y_j vary over C_j and requiring $y_i = x_i$ for $i \neq j$, (1) implies

$$f_j(y_j) + \sum_{i \neq j} f_i(x_i) \geq f_j(x_j) + \langle x_j^*, y_j - x_j\rangle + \sum_{i \neq j} (f_i(x_i) + \langle x_i^*, x_i - x_i\rangle).$$

Since all the numbers $f_i(x_i)$ are finite, this reduces to

$$f_j(y_j) \geq f_j(x_j) + \langle x_j^*, y_j - x_j\rangle,\ \forall y_j \in C_j. \tag{2}$$

But this is equivalent to $x_j^* \in \partial f_j(x_j)$. Thus we have shown that (1) implies $x_j^* \in \partial f_j(x_j)$ for $j = 1,\ldots,s$. The converse follows easily by summing the s inequalities of the form (2). This completes the proof of (e).

THEOREM 4.4. For $i = 1,\ldots,s$ let K_i be a closed proper concave-convex function on $R^{m_i} \times R^{n_i}$ with domain $C_i \times D_i$. Put $K = (K_1,\ldots,K_s)$ and write $C = C_1 \times\ldots\times C_s$, $D = D_1 \times\ldots\times D_s$, $x = (x_1,\ldots,x_s)$ and $y = (y_1,\ldots,y_s)$. Then the following statements hold:

(a) K is closed proper concave-convex with domain $C \times D$.

(b) If $\tilde{K}_i \in [K_i]$ for $i = 1,\ldots,s$, then $(\tilde{K}_1,\ldots,\tilde{K}_s) \in [K]$ (i.e. $[K]$ depends only on $[K_1],\ldots,[K_s]$).

(c) The least and greatest members of $[K]$ are given by

$$\underline{K}(x,y) = \begin{cases} \sum \underline{K}_i(x_i,y_i) & \text{if } x \in C \text{ and } y \in \operatorname{cl} D \\ +\infty & \text{if } x \in C \text{ and } y \notin \operatorname{cl} D \\ -\infty & \text{if } x \notin C \end{cases}$$

and

$$\overline{K}(x,y) = \begin{cases} \sum \overline{K}_i(x_i,y_i) & \text{if } x \in \operatorname{cl} C \text{ and } y \in D \\ -\infty & \text{if } x \notin \operatorname{cl} C \text{ and } y \in D \\ +\infty & \text{if } \qquad\qquad\quad y \notin D \end{cases}$$

(d) For $j = 1$ and 2 and $(x,y) \in C \times D$,

$$\partial_j K(x,y) = \partial_j K_1(x_1,y_1) \times\ldots\times \partial_j K_s(x_s,y_s)$$

(and $\partial K(x,y) = \emptyset$ whenever $(x,y) \notin C \times D$).

(e) $(K_1^*,\ldots K_s^*) \in [K^*]$

(f) $(\mathrm{rec}_1 K)(x) = \sum(\mathrm{rec}_1 K_i)(x_i)$ and $(\mathrm{rec}_2 K)(y) = \sum(\mathrm{rec}_2 K_i)(y_i)$

(g) If f (resp. f_i) denotes the convex parent of K (resp. K_i), then $f(x,y^*) = \sum f_i(x_i,y_i^*)$. Similarly for concave parents.

(h) If each K_i is polyhedral, then K is polyhedral.

PROOF. (a) It suffices to check that K satisfies the six conditions of (34.3). Let $x \in C$. Then $x_i \in C_i$ together with (34.3) applied to K_i imply that $K_i(x_i,\cdot)$ is a proper convex function with effective domain containing D_i. Since $K(x,y) = \sum K_i(x_i,y_i) + \delta(y|D)$, it follows from Lemma

4.3(a) and (5.2) that $K(x,\cdot)$ is proper convex with effective domain D. Now suppose $x \in \text{ri } C$. Then $x_i \in \text{ri } C_i$, so that (34.3) implies $K(x_i,\cdot)$ is closed and its effective domain actually equals D_i. Thus $K(x,y) = \sum K_i(x_i,y_i)$, and Lemma 4.3(b) implies $K(x,\cdot)$ is closed. This establishes the first two conditions of (34.3) for K. Of the remaining four conditions, two have parallel proofs and the other two are satisfied trivially.

(b) Let $\tilde{K}_i \in [K_i]$ for $i = 1,\ldots,s$ and write $\tilde{K} = (\tilde{K}_1,\ldots,\tilde{K}_s)$. Since by (34.4) two closed proper saddle-functions are equivalent iff they have the same kernel, $\tilde{K}_i(x_i,y_i) = K_i(x_i,y_i)$ whenever $(x_i,y_i) \in \text{ri } C_i \times \text{ri } D_i$. Hence $\tilde{K}$ and K agree on $\text{ri } C \times \text{ri } D$. Since equivalent saddle-functions have the same domain, $\text{dom } \tilde{K}_i = C_i \times D_i$. This implies $\text{dom } \tilde{K} = C \times D$. Therefore $\tilde{K}$ and K have the same kernel.

(c) Since K is closed, Theorem 0.1(b) implies $\underline{K} = \text{cl}_2 K$ and $\overline{K} = \text{cl}_1 K$. If $y \notin D$, then $K(\cdot,y)$ equals $+\infty$ on $\text{ri } C$ and hence $\overline{K}(\cdot,y) = \text{cl}(K(\cdot,y)) \equiv +\infty$. Now suppose $y \in D$. As in the proof of part (a), $K(x,y) = \sum K_i(x_i,y_i) - \delta(x|C)$ is proper concave with effective domain C. Since $g(x) = \sum K_i(x_i,y_i)$ is proper concave with $C \subset \text{dom } g \subset \text{cl } C$ by (34.3) and Lemma 4.3(a), it follows from (6.3.1) and (7.3.4) that $(\text{cl}_1 K)(x,y) = (\text{cl } g)(x)$. But $(\text{cl } g)(x) = \sum(\text{cl}_1 K_i)(x_i,y_i)$ by Lemma 4.3(b). Since $\text{cl}_1 K_i = \overline{K}_i$, this establishes the formula for $\overline{K}$. The other formula is proved similarly.

(d) By part (a) and (37.4), $\text{dom } \partial K \subset C \times D$. Suppose $(x,y) \in C \times D$. By (37.4.1), $\partial K(x,y) = \partial\overline{K}(\cdot,y)(x) \times \partial\underline{K}(x,\cdot)(y)$. But from part (c) and Lemma 4.3(e), $\partial\overline{K}(\cdot,y)(x) = \partial\overline{K}_1(\cdot,y_1)(x_1) \times\ldots\times \partial\overline{K}_s(\cdot,y_s)(x_s)$ where by (37.4.1) the $\overline{K}_i$ can be replaced by K_i. This establishes the assertion for $j = 1$, and the case $j = 2$ is exactly the same.

(e) The proof is by induction. First observe that separable saddle-functions can be given an equivalent, inductive definition. Namely, for

$s = 2$ let the definition be as given above, and for $s > 2$ define $(K_1,\ldots,K_s) = ((K_1,\ldots,K_{s-1}),K_s)$ where a space of the form $(R^{m_1} \times\ldots\times R^{m_{s-1}}) \times R^{m_s}$ is identified with $R^{m_1} \times\ldots\times R^{m_s}$. For the purpose of this proof we adopt this inductive definition. Suppose the assertion has already been proved for the case $s = 2$, and let $s > 2$ be fixed. Since $(K_1^*,\ldots,K_s^*) = ((K_1^*,\ldots,K_{s-1}^*),K_s^*)$ by definition, the inductive hypothesis

$$(K_1^*,\ldots,K_{s-1}^*) \in [(K_1,\ldots,K_{s-1})^*]]$$

together with parts (a) and (b) imply that $(K_1^*,\ldots,K_s^*)$ is equivalent to $((K_1,\ldots,K_{s-1})^*,K_s^*)$. But by the case $s = 2$ this is contained in $[((K_1,\ldots,K_{s-1}),K_s)^*]$, which by definition is the same as $[(K_1,\ldots,K_s)^*]$. Thus, part (e) will be proved once the case $s = 2$ is established. So let $s = 2$ and write $\mathrm{dom}\, K_1^* = C_1^* \times D_1^*$. By (36.3) and (36.1),

$$\underline{K}^*(x^*,y^*) = \sup_{y\in D}\inf_{x\in C} \{\Sigma\langle x_i,x_i^*\rangle + \langle y_i,y_i^*\rangle - K_i(x_i,y_i)\}$$

$$\leq \sup_{y_2\in D_2}\inf_{x_2\in C_2} \{\langle x_2,x_2^*\rangle + \langle y_2,y_2^*\rangle - K_2(x_2,y_2) + \underline{K}_1^*(x_1^*,y_1^*)\}$$

$$= \begin{cases} \Sigma\,\underline{K}_i^*(x_i^*,y_i^*) & \text{if } x_1^* \in C_1^* \text{ and } y_1^* \in \mathrm{dom}\,\underline{K}_1^*(x_1^*,\cdot) \\ +\infty & \text{if } x_1^* \in C_1^* \text{ and } y_1^* \notin \mathrm{dom}\,\underline{K}_1^*(x_1^*,\cdot) \\ -\infty & \text{if } x_1^* \notin C_1^* \end{cases}$$

Moreover, in the event that $x_1^* \in C_1^*$ and $y_1^* \in \mathrm{dom}\,\underline{K}_1^*(x_1^*,\ )$ we have

$$\Sigma\,\underline{K}^*_i(x_i^*,y_i^*) = \begin{cases} \Sigma\,\underline{K}_i^*(x_i^*,y_i^*) \in R & \text{if } x_2^* \in C_2^* \text{ and } y_2^* \in \mathrm{dom}\,\underline{K}_2^*(x_2^*,\ ) \\ +\infty & \text{if } x_2^* \in C_2^* \text{ and } y_2^* \notin \mathrm{dom}\,\underline{K}_2^*(x_2^*,\ ) \\ -\infty & \text{if } x_2^* \notin C_2^* \end{cases}$$

Also, $x_1^* \in C_1^*$ implies $D_1^* \subset \mathrm{dom}\,\underline{K}_1^*(x_1^*,\cdot)$ by (34.3). If $C^* = C_1^* \times C_2^*$ and $D^* = D_1^* \times D_2^*$, then the above facts imply

$$\mathrm{dom}_1\underline{K}^* \subset C^*,\quad D^* \subset \mathrm{dom}_2\underline{K}^*,$$

and

$$\underline{K}^*(x^*,y^*) \leq \sum \underline{K}_i^*(x_i^*,y_i^*) \quad \text{whenever} \quad x^* \in C^* \quad \text{or} \quad y^* \in D^*.$$

Parallel reasoning starting from $\overline{K}^*(x^*,y^*)$ yields that

$$C^* \subset \mathrm{dom}_1 \overline{K}^*, \quad \mathrm{dom}_2 \overline{K}^* \subset D^*,$$

and

$$\sum \overline{K}_i^*(x_i^*,y_i^*) \leq \overline{K}^*(x^*,y^*) \quad \text{whenever} \quad x^* \in C^* \quad \text{or} \quad y^* \in D^*.$$

Therefore $\mathrm{dom}\, K^* = C^* \times D^*$ and

$$\underline{K}^*(x^*,y^*) \leq \sum K_i^*(x_i^*,y_i^*) \leq \overline{K}^*(x^*,y^*)$$

whenever $x^* \in C^*$ or $y^* \in D^*$. By applying Theorem 0.1(b) to K^*, it follows that (K_1^*,K_2^*) and K^* have the same kernel. Since they are both closed and proper, (34.4) implies they are equivalent.

(f) From the definitions, Theorem 4.4(a), (34.3) and Lemma 4.3(f) it follows that

$$\begin{aligned}(\mathrm{rec}_2 K)(y) &= \sup\{(\mathrm{rec}\, K(x,\cdot))(y) \mid x \in \mathrm{ri}\, C\} \\ &= \sup\{\textstyle\sum (\mathrm{rec}\, K_i(x_i,\cdot))(y_i) \mid x_1 \in \mathrm{ri}\, C_1,\ldots,x_s \in \mathrm{ri}\, C_s\} \\ &= \textstyle\sum \sup\{(\mathrm{rec}\, K_i(x_i,\cdot))(y_i) \mid x_i \in \mathrm{ri}\, C_i\} \\ &= \textstyle\sum (\mathrm{rec}_2 K_i)(y_i).\end{aligned}$$

The other formula is proved similarly.

(g) By part (c), $\underline{K}(x,y) = \sum \underline{K}_i(x_i,y_i)$ whenever $x \in C$. Hence Theorem 0.1(a) implies that $f(x,y^*) = \sup\{\langle y,y^*\rangle - \sum \underline{K}_i(x_i,y_i)\} = \sum \sup_{y_i}\{\langle y_i,y_i^*\rangle - K_i(x_i,y_i)\} = \sum f_i(x_i,y_i^*)$ whenever $x \in C$. On the other hand, if $x \notin C$ then Theorem 0.1(a) implies that $f(x,\cdot)$ and $f_j(x_j,\cdot)$ for some $1 \leq j \leq s$ are constantly $+\infty$. Since each K_i is proper, each f_i is proper and hence never $-\infty$. Therefore $(x,y^*) \to \sum f_i(x_i,y_i^*)$ is constantly $+\infty$ whenever $x \notin C$. This establishes the formula.

(h) By part (g) and Lemma 4.3(c).

For the remainder of §4 let certain notation remain fixed as follows.

For $i = 1,\dots,s$ let K_i be a closed proper concave-convex function on $R^m \times R^n$ with domain $C_i \times D_i$. Write $C = C_1 \cap \dots \cap C_s$ and $D = D_1 \cap \dots \cap D_s$, and define $K = (K_1,\dots,K_s)$. Let A_1 map each $x \in R^m$ into the s-tuple $(x,\dots,x)$, let A_2 map each $y \in R^n$ into the s-tuple $(y,\dots,y)$, and put $A = A_1 \times A_2$.

The following condition is frequently used:

$$\mathrm{ri}(\mathrm{dom}\, K_1) \cap \dots \cap \mathrm{ri}(\mathrm{dom}\, K_s) \neq \emptyset \qquad (*)$$

The next lemma dualizes it.

LEMMA 4.5. <u>The condition</u> $\mathrm{ri}(\mathrm{dom}_1 K_1) \cap \dots \cap \mathrm{ri}(\mathrm{dom}_1 K_s) \neq \emptyset$ <u>is equivalent to</u>

$$\sum x_i^* = 0 \;\; \underline{\text{and}} \;\; \sum (\mathrm{rec}_1 K_i^*)(x_i^*) \geq 0 \Rightarrow \sum (\mathrm{rec}_1 K_i^*)(-x_i^*) \geq 0.$$

<u>Similarly, the condition</u> $\mathrm{ri}(\mathrm{dom}_2 K_1) \cap \dots \cap \mathrm{ri}(\mathrm{dom}_2 K_s) \neq \emptyset$ <u>is equivalent to</u>

$$\sum y_i^* = 0 \;\; \underline{\text{and}} \;\; \sum (\mathrm{rec}_2 K_i^*)(y_i^*) \leq 0 \Rightarrow \sum (\mathrm{rec}_2 K_i^*)(-y_i^*) \leq 0.$$

PROOF. Apply Lemma 0.5 to the saddle-function $(K_1^*,\dots,K_s^*)$ and the subspace $\{(x_1^*,\dots,x_s^*) \mid x_1^* = \dots = x_s^*\}$ and simplify using Theorem 4.4. The second assertion is proved similarly.

The next theorem enables us to apply the results of §§2 and 3 to the equivalence class $[K_1 + \dots + K_s]$.

THEOREM 4.6. <u>Assume</u> $(*)$. <u>Then</u> $[K_1 + \dots + K_s]$ <u>is defined and equals</u> $[KA]$.

PROOF. Theorem 4.2 implies $[K_1 + \dots + K_s]$ is defined and has kernel

$$(x,y) \to \sum K_i(x,y), \; \forall (x,y) \in \mathrm{ri}\, C \times \mathrm{ri}\, D.$$

Theorems 4.4(a) and 2.2 imply that $[KA]$ exists, and it is easy to check that its kernel is the function given above. The theorem now follows from (34.4).

COROLLARY 4.6.1. <u>Assume</u> $(*)$. <u>If each</u> $[K_i]$ <u>is polyhedral, then</u> $[K_1 + \dots + K_s]$ <u>is polyhedral.</u>

PROOF. By Theorem 4.4(h) and Corollary 2.4.1.

COROLLARY 4.6.2. *Assume* (✱). *If* h (*resp.* k) *denotes the convex* (*resp. concave*) *parent of* $[K_1 + \ldots + K_s]$ *and* f_i (*resp.* g_i) *denotes the convex* (*resp. concave*) *parent of* $[K_i]$, *then*

$$h(x,y^*) = \inf\{\sum f_i(x,y_i^*) \mid \sum y_i^* = y^*\}$$

and

$$k(x^*,y) = \sup\{\sum g_i(x_i^*,y) \mid \sum x_i^* = x^*\}.$$

PROOF. By Theorems 2.4 and 4.4(g).

COROLLARY 4.6.3. *Assume* (✱). *Then*

$$\operatorname{dom}(K_1 + \ldots + K_s)^* \subset \operatorname{dom} K_1^* + \ldots + \operatorname{dom} K_s^*.$$

In particular, if f_i *and* g_i *are as in Corollary 4.6.2, then*

$$\operatorname{ri}(\operatorname{dom}_1(K_1 + \ldots + K_s)^*) = \bigcup\{\sum \operatorname{ri}(\operatorname{dom} g_i(\cdot,y)) \mid y \in \operatorname{ri} D\}$$

and

$$\operatorname{ri}(\operatorname{dom}_2(K_1 + \ldots + K_s)^*) = \bigcup\{\sum \operatorname{ri}(\operatorname{dom} f_i(x,\cdot)) \mid x \in \operatorname{ri} C\}$$

where these formulas also hold with "ri" *deleted throughout.*

PROOF. Apply Corollary 2.4.2, using Theorem 4.4(g) and Lemma 4.3(a) to simplify.

In convex function theory there is a result correspondin^ to the inclusion in Corollary 4.6.3. Namely,

$$\operatorname{dom}(f_1 + \ldots + f_s)^* = \operatorname{dom} f_1^* + \ldots + \operatorname{dom} f_s^*$$

whenever $f_1,\ldots,f_s$ are proper convex functions satisfying $\operatorname{ri}(\operatorname{dom} f_1)\cap\ldots\cap\operatorname{ri}(\operatorname{dom} f_s) \neq \emptyset$ (see(16.4)). One might hope in the saddle-function case to have at least

$$\operatorname{ri}(\operatorname{dom} K_1^* + \ldots + \operatorname{dom} K_s^*) \subset \operatorname{dom}(K_1 + \ldots + K_s)^* \subset \operatorname{dom} K_1^* + \ldots + \operatorname{dom} K_s^*$$

satisfied whenever $K_1,\ldots,K_s$ are closed proper concave-convex functions satisfying $\operatorname{ri}\operatorname{dom} K_1\cap\ldots\cap\operatorname{ri}\operatorname{dom} K_s \neq \emptyset$. However this can fail drastically, as can be seen by taking $s = 2$ and putting $K_1(x,y) = \langle x,y\rangle$ and

$K_2(x,y) = -\langle x,y\rangle$ on $R^m \times R^m$. In this case $\operatorname{dom}(K_1 + K_2)^* = \{0\} \times \{0\}$ whereas $\operatorname{dom} K_1^* = R^m \times R^m = \operatorname{dom} K_2^*$. Theorems 4.8 and 4.11 give conditions which guarantee that this "collapsing" behavior of $\operatorname{dom}(K_1 +\ldots+ K_s)^*$ cannot occur.

LEMMA 4.7. <u>Assume</u> $(\ast)$. <u>Then</u>

$$(\mathrm{rec}_1(K_1 +\ldots+ K_s))(x) = \inf\{\Sigma(\mathrm{rec}\, K_i(\cdot,y))(x) \mid y \in \mathrm{ri}\, D\}$$

<u>and</u>

$$(\mathrm{rec}_2(K_1 +\ldots+ K_s))(y) = \sup\{\Sigma(\mathrm{rec}\, K_i(x,\cdot))(y) \mid x \in \mathrm{ri}\, C\}.$$

PROOF. By Theorem 4.6 the formulas in Lemma 2.5 can be applied. Simplify using (34.3) and Lemma 4.3(f).

THEOREM 4.8. <u>Assume</u> $(\ast)$. <u>Then for</u> $j = 1$ <u>and</u> 2,

$$\mathrm{cl}(\mathrm{dom}_j(K_1 +\ldots+ K_s)^*) = \mathrm{cl}(\mathrm{dom}_j K_1^* +\ldots+ \mathrm{dom}_j K^*_s)$$

<u>iff</u>

$$\mathrm{rec}_j(K_1 +\ldots+ K_s) = \mathrm{rec}_j K_1 +\ldots+ \mathrm{rec}_j K_s.$$

PROOF. By Theorems 4.6 and 2.6.

The next theorem parallels the result obtained by Rockafellar [38], Moreau [32], and others for the subdifferential of a sum of convex functions.

THEOREM 4.9. <u>Assume</u> $(\ast)$. <u>Then</u>

$$\partial(K_1 +\ldots+ K_s)(x,y) = \partial K_1(x,y) +\ldots+ \partial K_s(x,y)$$

<u>for each</u> $(x,y) \in R^m \times R^n$.

PROOF. Since $\operatorname{dom}(K_1 +\ldots+ K_s) = C \times D = \operatorname{dom} K_1 \cap\ldots\cap \operatorname{dom} K_s$, (37.4) implies that $\partial(K_1 +\ldots+ K_s)(x,y)$ and $\partial K_j(x,y)$ are empty (for some j) whenever $(x,y) \notin C \times D$. So suppose that $(x,y) \in C \times D$. Then Theorems 4.6 and 2.3 imply that

$$\partial(K_1 +\ldots+ K_s)(x,y) = A^*\partial K(A(x,y)).$$

The formula follows from this together with Theorem 4.4(d) and the definitions, after observing that A_1^* and A_2^* are just the appropriate addition

linear transformations.

The next theorem identifies certain members of the equivalence class conjugate to $[K_1 + \ldots + K_s]$.

THEOREM 4.10. *Assume* $(\ast)$. *Let* $\operatorname{dom} K_i^* = C_i^* \times D_i^*$, *and define functions* ϕ *and* ψ *on* $R^m \times R^n$ *by*

$$\phi(z,w) = \sup_{\substack{\Sigma z_i = z \\ z_i \in C_i^*}} \inf_{\Sigma w_i = w} \Sigma \underline{K}^*_i(z_i,w_i)$$

and

$$\psi(z,w) = \inf_{\substack{\Sigma w_i = w \\ w_i \in D_i^*}} \sup_{\Sigma z_i = z} \Sigma \overline{K}_i^*(z_i,w_i)$$

Then $[(K_1 + \ldots + K_s)^*]$ *contains each concave-convex function* J *on* $R^m \times R^n$ *satisfying* $\phi \le J \le \psi$. *If each* K_i *is polyhedral, then* $[(K_1 + \ldots + K_s)^*]$ *is polyhedral.*

PROOF. By Theorems 4.6 and 2.8, $[(K_1 + \ldots + K_s)^*]$ contains each saddle-function lying between two certain functions J_1 and J_2. By parts (e) and (c) of Theorem 4.4, one can easily show that $J_1 = \phi$ and $J_2 = \psi$. The polyhedral assertion is immediate from Corollary 4.6.1 and the fact that $[K^*]$ is polyhedral whenever $[K]$ is.

There are actually many representations of the functions ϕ and ψ in Theorem 4.10. Suppose Z_i and W_i are any sets (not necessarily convex) such that $\operatorname{ri} C_i^* \subset Z_i \subset R^m$ and $\operatorname{ri} D_i^* \subset W_i \subset R^n$ (for $i = 1,\ldots,s$). Then

$$\phi(z,w) = \sup_{\substack{\Sigma z_i = z \\ z_i \in C_i^*}} \inf_{\substack{\Sigma w_i = w \\ w_i \in W_i}} \Sigma \underline{K}_i^*(z_i,w_i)$$

and

$$\psi(z,w) = \inf_{\substack{\Sigma w_i = w \\ w_i \in D_i^*}} \sup_{\substack{\Sigma z_i = z \\ z_i \in Z_i}} \Sigma \overline{K}_i^*(z_i,w_i).$$

This follows from the observation that, if $A: R^n \to R^m$ is a linear transformation and f is a proper convex function on R^n, then $(Af)(y) = \inf\{f(x) \mid Ax = y, x \in C\}$ for any set C such that $\mathrm{ri}(\mathrm{dom}\, f) \subset C \subset R^n$.

The fact that $[(K_1 + \ldots + K_s)^*]$ contains ϕ and ψ suggests writing

$$[(K_1 + \ldots + K_s)^*] = [K_1^*] \square \ldots \square [K_s^*]$$

and calling this class the minimax convolution of $[K_1^*], \ldots, [K_s^*]$. This is the saddle-function analogue of the operation of infimal convolution on convex functions. The identity above expresses the fact that the operations of addition and minimax convolution are dual, just as in convex function theory the formula $(f_1 + \ldots + f_s)^* = f_1^* \square \ldots \square f_s^*$ expresses the duality between the operations of addition and infimal convolution.

The next theorem gives information concerning attainment of the extrema appearing in the definitions of ϕ and ψ.

THEOREM 4.11. *Let ϕ and ψ be defined as in Theorem 4.10, and assume that whenever $(z_i, w_i) \in \mathrm{ri}(\mathrm{dom}\, K_i^*)$ for $i = 1, \ldots, s$ the following conditions are satisfied:*

(a) $\sum \bar{w}_i = 0$ *and* $\sum (\mathrm{rec}\, \underline{K}_i^*(z_i, \cdot))(\bar{w}_i) \le 0 \Rightarrow \sum (\mathrm{rec}\, \underline{K}_i^*(z_i, \cdot))(-\bar{w}_i) \le 0$

(b) $\sum \bar{z}_i = 0$ *and* $\sum (\mathrm{rec}\, \overline{K}_i^*(\cdot, w_i))(\bar{z}_i) \ge 0 \Rightarrow \sum (\mathrm{rec}\, \overline{K}_i^*(\cdot, w_i))(-\bar{z}_i) \ge 0$.

Then the conclusions of Theorem 4.10 hold and

$$\mathrm{ri}(\mathrm{dom}\, K_1^* + \ldots + \mathrm{dom}\, K_s^*) \subset \mathrm{dom}\, (K_1 + \ldots + K_s)^* \subset \mathrm{dom}\, K_1^* + \ldots + \mathrm{dom}\, K_s^*.$$

Moreover, for each $(z,w) \in \mathrm{ri}(\mathrm{dom}\, K_1^* + \ldots + \mathrm{dom}\, K_s^*)$ there exist nonempty closed convex sets $Z \subset R^{sm}$ and $W \subset R^{sn}$ such that for each $(\bar{z}_1, \ldots, \bar{z}_s) \in Z$ and $(\bar{w}_1, \ldots, \bar{w}_s) \in W$ the following statements hold:

(i) $\sum (\bar{z}_i, \bar{w}_i) = (z,w)$ *and* $(\bar{z}_i, \bar{w}_i) \in \mathrm{dom}\, \partial K_i^*$ *for each* i;

(ii) $\phi(z,w) = \sum K_i^*(\bar{z}_i, \bar{w}_i) = \psi(z,w)$;

(iii) $\sum K_i^*(z_i, \bar{w}_i) \le \sum K_i^*(\bar{z}_i, \bar{w}_i) \le \sum K_i^*(\bar{z}_i, w_i)$ *whenever* $\sum (z_i, w_i) = (z,w)$ *and* $(z_i, w_i) \in \mathrm{dom}\, K_i^*$ *for each* i.

PROOF. By parts (e) and (c) of Theorem 4.4 together with Lemma 4.3(f), A^* and $K^* = (K_1^*,\ldots,K_s^*)$ satisfy the hypotheses of Theorem 3.4. The assertions are immediate from this and Theorem 4.4(d).

If conditions (a) and (b) above are actually satisfied whenever $(z_i,w_i) \in \operatorname{dom} K_i^*$ for $i = 1,\ldots,s$, then Theorem 3.5 implies that $\operatorname{dom}(K_1 +\ldots+ K_s)^* = \operatorname{dom} K_1^* +\ldots+ \operatorname{dom} K_s^*$ and that ϕ and ψ are the least and greatest members of $[(K_1 +\ldots+ K_s)^*]$.

The following lemma may be useful in applying Theorem 4.11. Notice for example that its conditions are satisfied when each of the sets $\operatorname{dom} K_i^*$ is bounded.

LEMMA 4.12. <u>The following condition implies that condition (a) of Theorem 4.11 is satisfied for each choice of</u> $z_1 \in \operatorname{dom}_1 K_1^*,\ldots,z_s \in \operatorname{dom}_1 K_s^*$:

(c) $\sum w_i = 0$ <u>and</u> $w_i \in 0^+ \operatorname{cl}(\operatorname{dom}_2 K_i^*)$ <u>for each</u> i <u>imply that</u> $w_i = 0$ <u>for each</u> i.

<u>Similarly, the following condition implies that condition (b) of Theorem 4.11 is satisfied for each choice of</u> $w_1 \in \operatorname{dom}_2 K_1^*,\ldots,w_s \in \operatorname{dom}_2 K_s^*$:

(d) $\sum z_i = 0$ <u>and</u> $z_i \in 0^+ \operatorname{cl}(\operatorname{dom}_1 K_i^*)$ <u>for each</u> i <u>imply that</u> $z_i = 0$ <u>for each</u> i.

PROOF. Apply Lemma 3.6 to A^* and $K^* = (K_1^*,\ldots,K_s^*)$. Condition (c) (resp. (d)) corresponds to condition (c_1) (resp. (d_1)) of Lemma 3.6, and condition (a) (resp. (b)) of Theorem 4.11 corresponds to condition (a_3) (resp. (b_3)) of Lemma 3.3.

The next lemma furnishes alternate characterizations of conditions (c) and (d) of Lemma 4.12

LEMMA 4.13. <u>Let</u> $P_1,\ldots,P_s$ <u>be convex cones in</u> R^n <u>which contain the origin. Then the following conditions are equivalent:</u>

(i) $\sum p_i = 0$ <u>and</u> $p_i \in P_i$ <u>for each</u> i <u>imply</u> $p_i = 0$ <u>for each</u> i;

(ii) $(-P_j)\cap(\text{conv}\bigcup_{i\neq j}P_i) = \{0\}$ for each $j = 1,\ldots,s.$

PROOF. First, observe that for each j, (3.3) implies

$$\text{conv}\bigcup_{i\neq j}P_i = \bigcup\{\sum_{i\neq j}\lambda_i P_i \mid 0 \leq \lambda_i \text{ and } \sum\lambda_i = 1\}.$$

From this it follows that $\text{conv}\bigcup_{i\neq j}P_i = \sum_{i\neq j}P_i$. Thus, condition (ii) fails iff

$$\exists_j \text{ and } \exists p_j \in P_j \text{ such that } 0 \neq -p_j \in \sum_{i\neq j}P_i.$$

This occurs iff

$$\exists p_1 \in P_1,\ldots,\exists p_s \in P_s \text{ and } \exists_j \text{ such that } 0 \neq -p_j = \sum_{i\neq j}p_i,$$

which occurs iff condition (i) fails.

We conclude this section with an example concerning maximal monotone operators arising from saddle-functions. This will suggest a conjecture about arbitrary maximal monotone operators.

By (37.5.2), each closed proper concave-convex function K on $R^m \times R^n$ induces a maximal monotone operator T (generally multivalued) from $R^m \times R^n$ to $R^m \times R^n$ by means of the formula

$$T(x,y) = \{(-x^*,y^*) \mid (x^*,y^*) \in \partial K(x,y)\}.$$

By (37.4.1), T depends only on the equivalence class containing K. If $R(\cdot)$ denotes the range of an operator and B is the linear transformation which sends (x^*,y^*) to $(-x^*,y^*)$, then (37.5) implies that

$$R(T) = B\ \text{dom}\ \partial K$$

whenever T arises from K as above.

EXAMPLE 4.14. Assume that conditions (c) and (d) of Lemma 4.12 are satisfied. Then by Lemma 4.12 the hypotheses of Theorem 4.11 are met, and these in turn imply that condition $(\ast)$ is satisfied. Let B be the linear transformation defined above, let T_i be the maximal monotone operator induced by K_i as described above, and similarly (using Theorem 4.5) let

T be the maximal monotone operator induced by ΣK_i. By (37.4), (6.3.1) and (9.1) it follows that $\text{cl } R(T_i) = B \text{ cl dom } K_i^*$, and similarly $\text{cl } R(T) = B \text{ cl dom}(\Sigma K_i)^*$. Theorem 4.11 and (6.3.1) imply $\text{cl dom}(\Sigma K_i)^* = \text{cl } \Sigma \text{ dom } K_i^*$. Combining these facts with (6.6.2) yields $\Sigma \text{ cl } R(T_i) \subset \text{cl } R(T)$. Since Theorem 4.7 implies $\Sigma T_i = T$, this shows that ΣT_i is a maximal monotone operator satisfying

$$\Sigma \text{ cl } R(T_i) \subset \text{cl } R(\Sigma T_i). \tag{1}$$

Furthermore, it can be deduced from (1), $R(\Sigma T_i) \subset \Sigma R(T_i)$ and (9.1) that

$$\Sigma\, 0^+ \text{ cl } R(T_i) \subset 0^+ \text{ cl } R(\Sigma T_i).$$

It is easy to show that conditions (c) and (d) of Lemma 4.12 can be reformulated equivalently as follows:

$$\left.\begin{array}{c} \Sigma z_i = 0 \text{ and } z_i \in 0^+ \text{ cl } R(T_i) \\ \text{for each } i \text{ implies } z_i = 0 \text{ for each } i \end{array}\right\} \tag{++}$$

Now (++) and (9.1.1) imply $\Sigma \text{ cl } R(T_i) = \text{cl } \Sigma R(T_i)$, $R(\Sigma T_i) \subset \Sigma R(T_i) \subset \Sigma \text{ cl } R(T_i)$ holds trivially. Combining these facts with (1) yields

$$\Sigma \text{ cl } R(T_i) = \text{cl } R(\Sigma T_i). \tag{2}$$

Furthermore, from (2), (++) and (9.1.1) it follows immediately that

$$\Sigma\, 0^+ \text{ cl } R(T_i) = 0^+ \text{ cl } R(\Sigma T_i).$$

It is known that ΣT_i is a maximal monotone operator satisfying (1) whenever each T_i is the subdifferential of a closed proper convex function on R^n and the condition

$$\text{ri } D(T_1) \cap \ldots \cap \text{ri } D(T_s) \neq \emptyset \tag{+}$$

is satisfied, where $D(T) = \{z \mid T(z) \neq \emptyset\}$. Moreover, in this situation (2) actually holds if (++) is satisfied.

On the other hand, (1) fails in general for maximal monotone operators satisfying (+). (For example, take $s = 2$ and consider the T_i's induced

by the saddle-functions K_1 and K_2 defined following Corollary 4.6.3.) It is not known, though, whether (2) holds for arbitrary maximal monotone operators satisfying (++). But the fact that this formula does hold for those operators arising from saddle-functions leads one to conjecture that it holds in general. This is because such operators, unlike the subdifferentials of convex functions, exhibit most of the pathology of arbitrary maximal monotone operators. Indeed, the last fact is one of the main motivations for studying saddle-functions.

§5: The Partial Conjugacy Operation

In this section the results of §2 are used to develop another operation on equivalence classes of closed proper saddle-functions. It follows from Theorems 5.1 and 5.2 that this operation induces a symmetric one-to-one correspondence among such equivalence classes. In §6 this operation is used to assign a well-defined Lagrangian to each dual pair of generalized saddle programs.

Throughout §5 let K be a closed proper concave-convex function on $(R^p \times R^m) \times (R^q \times R^n)$, and let W_1 and W_2 be functions on $(R^p \times R^n) \times (R^q \times R^m)$ defined by

$$W_1(u^*,y,v^*,x) = \sup_v \inf_u \{\langle u^*,u\rangle + \langle v^*,v\rangle - K(u,x,v,y)\}$$

and

$$W_2(u^*,y,v^*,x) = \inf_u \sup_v \{\langle u^*,u\rangle + \langle v^*,v\rangle - K(u,x,v,y)\}.$$

THEOREM 5.1. <u>The functions</u> W_1 <u>and</u> W_2 <u>belong to an equivalence class</u> $[W]$ <u>of closed proper concave-convex functions. Furthermore,</u> $[W]$ <u>depends only on</u> $[K]$, <u>and</u> $[W]$ <u>is polyhedral if</u> $[K]$ <u>is</u>.

PROOF. Define a linear transformation $A = A_1 \times A_2$ and a function H by

$$A_1(v,u^*,y) = (u^*,y),$$
$$A_2(u,v^*,x) = (v^*,x),$$

and

$$H(v,u^*,y,u,v^*,x) = \langle u,u^*\rangle + \langle v,v^*\rangle - K(u,x,v,y).$$

Clearly H is closed proper concave-convex on $(R^q \times R^p \times R^n) \times (R^p \times R^q \times R^m)$. If $(v,y) \in \mathrm{ri}(\mathrm{dom}_2 K)$, (34.3) implies that the function

$$(u,v^*,x) \to -K(u,x,v,y)$$

is closed proper convex, and by (8.5) its recession function can be shown to be

$$(u,v^*,x) \to -(\mathrm{rec}\, K(\cdot,\cdot,v,y))(u,x).$$

Also, the function

$$(u,v^*,x) \to \langle u,u^*\rangle + \langle v,v^*\rangle$$

is closed proper convex and coincides with its recession function. Hence (9.3) implies that

$$(\text{rec } H(v,u^*,y,\cdot,\cdot,\cdot))(u,v^*,x) = \langle u,u^*\rangle + \langle v,v^*\rangle - (\text{rec } K(\cdot,\cdot,v,y))(u,x)$$

whenever $(v,y) \in \text{ri}(\text{dom}_2 K)$. Therefore $A_2^{-1}\{0\} \cap (\text{rec cone}_2 H)$ equals

$$\{(u,0,0) \mid \langle u,u^*\rangle - (\text{rec } K(\cdot,\cdot,v,y))(u,0) \leq 0, \forall u^* \in R^p, \forall (v,y) \in \text{ri}(\text{dom}_2 K)\}.$$

Now by (34.3) and (8.5), $(v,y) \in \text{ri}(\text{dom}_2 K)$ implies that $\text{rec } K(\cdot,\cdot,v,y)$ is never $+\infty$. It follows that $A_2^{-1}\{0\} \cap (\text{rec cone}_2 H)$ is the nullspace of $R^p \times R^q \times R^m$. Similarly, $A_1^{-1}\{0\} \cap (\text{rec cone}_1 H)$ is the nullspace of $R^q \times R^p \times R^n$. Therefore by Lemma 2.9, range $A^* \cap \text{ri}(\text{dom } H^*) \neq \emptyset$. The first two assertions of the theorem now follow from Theorem 2.8 and the fact that $\underline{K} \leq \tilde{K} \leq \overline{K}$ whenever $\tilde{K} \in [K]$. If K is polyhedral, then Corollary 4.6.1 implies H is polyhedral and hence Theorem 2.8 implies $[AH] = [W]$ is polyhedral.

The equivalence class $[W]$ containing W_1 and W_2 is called the <u>partial conjugate of</u> $[K]$ <u>in</u> u <u>and</u> v, and the operation which sends $[K]$ to $[W]$ is called <u>partial conjugacy</u>. This terminology is suggested by the fact that forming $[W]$ involves only parts of the arguments of K, whereas forming the (ordinary) conjugate $[K^*]$ involves all of the arguments of K.

THEOREM 5.2 <u>The partial conjugate of</u> $[W]$ <u>in</u> u^* <u>and</u> v^* <u>is</u> $[K]$.

PROOF. By Theorem 5.1, $[W]$ contains the function $\tilde{W}$, where $\tilde{W}(u^*,y,v^*,x) = \inf_{\bar{u}} \sup_{\bar{v}} (\langle \bar{u},u^*\rangle + \langle \bar{v},v^*\rangle - \underline{K}(\bar{u},x,\bar{v},y))$. Hence the partial conjugate of $[W]$ in u^* and v^* contains the function M given by

$$\begin{aligned} M(u,x,v,y) &= \sup_{v^*} \inf_{u^*} (\langle u,u^*\rangle + \langle v,v^*\rangle - \tilde{W}(u^*,y,v^*,x)) \\ &= \sup_{v^*} \inf_{u^*} \sup_{\bar{u}} \inf_{\bar{v}} (\langle u^*,u-\bar{u}\rangle + \langle v^*,v-\bar{v}\rangle + \underline{K}(\bar{u},x,\bar{v},y)). \end{aligned}$$

By the same technique used in the proof of Theorem 5.1 it can be verified

that range $B^* \cap \mathrm{ri}(\mathrm{dom}\, J^*) \neq \emptyset$, where $B = B_1 \times B_2$ and J are given by

$$B_1(v^*,\bar{u},u,x) = (u,x),$$

$$B_2(u^*,\bar{v},v,y) = (v,y),$$

and

$$J(v^*,\bar{u},u,x,u^*,\bar{v},v,y) = \langle u^*,u - \bar{u}\rangle + \langle v^*, v - \bar{v}\rangle + \underline{K}(\bar{u},x,\bar{v},y).$$

Therefore Theorem 2.8 implies that [BJ] is well-defined. Now by (36.1) and Theorem 0.1(b) it follows easily that M and N belong to [BJ], where N is given by

$$N(u,x,v,y) = \sup_{\bar{u}}\ \inf_{\bar{v}}\ \sup_{v^*}\ \inf_{u^*}\ \{\langle u^*,u - \bar{u}\rangle + \langle v^*,v - \bar{v}\rangle + \underline{K}(\bar{u},x,\bar{v},y)\}.$$

Thus, to complete the proof it suffices to show that $N \in [K]$.

Let u,x,v,y be arbitrary but fixed. For each $\bar{u}$ define

$$p(\bar{u}) = \inf_{\bar{v}}\ \sup_{v^*}\ \inf_{u^*}\ \{\langle u^*,u - \bar{u}\rangle + \langle v^*,v - \bar{v}\rangle + \underline{K}(\bar{u},x,\bar{v},y)\}.$$

Observe that

$$N(u,x,v,y) = \sup\{\, p(\bar{u}) \mid \bar{u} \in U\,\}, \tag{1}$$

where $U = \{\bar{u} \mid (\bar{u},x) \in \mathrm{dom}_1 K\}$. Indeed, if $(\bar{u},x) \notin \mathrm{dom}_1 K$ then $\underline{K}(\bar{u},x,\cdot,\cdot) \equiv -\infty$ so that $p(\bar{u}) = -\infty$. Thus,

$$N(u,x,v,y) = -\infty = \underline{K}(u,x,v,y) \tag{2}$$

whenever $U = \emptyset$. Now assume $U \neq \emptyset$. For each $\bar{u} \in U$, $\underline{K}(\bar{u},x,\cdot,\cdot)$ is never $-\infty$ and hence

$$p(\bar{u}) = \inf_{\bar{v} \in V(\bar{u})}\ \sup_{v^*}\ \inf_{u^*}\ \{\langle u^*,u - \bar{u}\rangle + \langle v^*,v - \bar{v}\rangle + \underline{K}(\bar{u},x,\bar{v},y)\},$$

where $V(\bar{u}) = \{\bar{v} \mid \underline{K}(\bar{u},x,\bar{v},y) < +\infty\}$. This implies $p(\bar{u}) = +\infty$ whenever $V(\bar{u}) = \emptyset$. Hence (1) implies $N(u,x,v,y) = +\infty$ if there exists a $\bar{u} \in U$ such that $V(\bar{u}) = \emptyset$. But for such a $\bar{u}$, $\underline{K}(\bar{u},x,v,y) = +\infty$. Therefore

$$N(u,x,v,y) = +\infty = \overline{K}(u,x,v,y) \tag{3}$$

whenever there exists a $\bar{u} \in U$ such that $V(\bar{u}) = \emptyset$. Finally, assume $U \neq \emptyset$ and $V(\bar{u}) \neq \emptyset$ for every $\bar{u} \in U$. Then for each $\bar{u} \in U$,

$$p(\bar{u}) = \inf_{\bar{v}\in V(\bar{u})} \{\underline{K}(\bar{u},x,\bar{v},y) + \sup_{v^*} \{\langle v^*, v - \bar{v}\rangle + \inf_{u^*} \{\langle u^*, u - \bar{u}\rangle\}\}\}$$

$$= \begin{cases} -\infty & \text{if } \bar{u} \neq u \\ \inf_{\bar{v}\in V(u)} \{\underline{K}(u,x,\bar{v},y) + \sup_{v^*} \{\langle v^*, v - \bar{v}\rangle\}\} & \text{if } \bar{u} = u, \end{cases}$$

where

$$p(u) = \inf \{\underline{K}(u,x,\bar{v},y) \mid \bar{v} \in V(u),\ \bar{v} = v\}$$

$$= \begin{cases} +\infty & \text{if } v \notin V(u) \\ \underline{K}(u,x,v,y) & \text{if } v \in V(u). \end{cases}$$

Hence (1) implies that in this case

$$N(u,x,v,y) = \sup \{p(\bar{u}) \mid \bar{u} \in U,\ \bar{u} = u\}$$

$$= \begin{cases} -\infty & \text{if } u \notin U \\ +\infty & \text{if } u \in U \text{ and } v \notin V(u) \\ \underline{K}(u,x,v,y) & \text{if } u \in U \text{ and } v \in V(u) \end{cases}$$

$$= \underline{K}(u,x,v,y).$$

Combining this with (2) and (3) yields $\underline{K} \leq N \leq \overline{K}$ everywhere. Hence Theorem 0.1(b) implies $N \in [K]$.

THEOREM 5.3. The following conditions are equivalent:

(a) $(u^*,x^*,v^*,y^*) \in \partial K(u,x,v,y)$

(b) $(u,-y^*,v,-x^*) \in \partial W(u^*,y,v^*,x)$

PROOF. By (37.5) condition (b) is equivalent to

$$(u^*,y,v^*,x) \in \partial W^*(u,-y^*,v,-x^*).$$

But from the proof of Theorem 5.1 we know that $[W^*] = [H^*A^*]$ and range $A^* \cap \mathrm{ri}(\mathrm{dom}\, H^*) \neq \emptyset$. Hence by Theorem 2.3,

$$\partial W^*(u,-y^*,v,-x^*) = A\partial H^*(A^*(u,-y^*,v,-x^*)).$$

It follows that condition (b) is equivalent to the existence of points u^0 and v^0 such that

$$(v^0,u^*,y,u^0,v^*,x) \in \partial H^*(o,u,-y^*,o,v,-x^*).$$

But by (37.5) and (37.4) this containment occurs iff (v^0,u^*,y,u^0,v^*,x) is a saddle-point of

$$H - \langle\cdot,(o,u,-y^*)\rangle - \langle\cdot,(o,v,-x^*)\rangle$$

and $H(v^0,u^*,y,u^0,v^*,x) \in R$. Therefore by the definition of H, condition (b) is equivalent to the existence of points u^0 and v^0 such that $K(u^0,x,v^0,y) \in R$ and

$$\begin{aligned} &K(\tilde{u},\tilde{x},v^0,y) - \langle\tilde{u} - u,u^*\rangle - \langle v^0 - v,\tilde{v}^*\rangle - \langle\tilde{x} - x,x^*\rangle \\ &\leq K(u^0,x,v^0,y) - \langle u^0 - u,u^*\rangle - \langle v^0 - v,v^*\rangle \\ &\leq K(u^0,x,\tilde{v},\tilde{y}) - \langle u^0 - u,\tilde{u}^*\rangle - \langle\tilde{v} - v,v^*\rangle - \langle\tilde{y} - y,y^*\rangle \end{aligned}$$

for all $(\tilde{v},\tilde{u}^*,\tilde{y})$ and $(\tilde{u},\tilde{v}^*,\tilde{x})$. Now pick any $(v',y') \in \mathrm{dom}_2K$. Choosing $\tilde{v} = v'$ and $\tilde{y} = y'$ in the above condition implies

$$K(u^0,x,v',y') \geq \alpha + \langle u^0 - u,\tilde{u}^*\rangle \text{ for all } \tilde{u}^*,$$

where α is a certain real constant. Thus if u^0 were different from u, we would have $K(u^0,x,v',y') = +\infty$, contradicting $(v',y') \in \mathrm{dom}_2K$. Hence in the above condition we must have $u^0 = u$, and similarly $v^0 = v$. Therefore condition (b) is equivalent to $(K(u,x,v,y) \in R$ and)

$$K(\tilde{u},\tilde{x},v,y) - \langle\tilde{u} - u,u^*\rangle - \langle\tilde{x} - x,x^*\rangle \leq K(u,x,v,y), \ \forall(\tilde{u},\tilde{x})$$

and

$$K(u,x,v,y) \leq K(u,x,\tilde{v},\tilde{y}) - \langle\tilde{v} - v,v^*\rangle - \langle\tilde{y} - y,y^*\rangle, \ \forall(\tilde{v},\tilde{y}).$$

But these conditions are clearly equivalent to (a).

§6: Generalized Saddle Programs

In this section the results of §§2, 3 and 5 are applied to the problem of associating with a given minimax problem a dual minimax problem and developing a perturbational duality theory for such pairs of problems.

Ignoring technical details, we can outline the general approach as follows. Suppose we are given a minimax problem in the form of an equivalence class $[K_0]$ of saddle-functions on $R^m \times R^n$. This minimax problem is first extended to a saddle program in the form of another equivalence class $[K]$ of saddle-functions on $(R^p \times R^m) \times (R^q \times R^n)$. The extension is such that $[K_0]$ is suitably "embedded" in $[K]$, i.e., the saddle-functions $(x,y) \to \tilde{K}(o,x,o,y)$ for $\tilde{K} \in [K]$ are all required to belong to $[K_0]$. By a modification of the conjugacy correspondence, an equivalence class $[L]$ of saddle-functions on $(R^m \times R^p) \times (R^n \times R^q)$ is then obtained from $[K]$. The saddle program given by $[L]$ is called the dual of the program given by $[K]$. Under a mild hypothesis on $[K]$, the saddle-functions $(z,w) \to \tilde{L}(o,z,o,w)$ for $\tilde{L} \in [L]$ all belong to a single equivalence class $[L_0]$. In this event the minimax problem given by $[L_0]$ is the dual of the minimax problem given by $[K_0]$. In this sense $[K_0]$ may have many such duals, since $[L]$ and hence $[L_0]$ depends not only on $[K_0]$ but also on the particular "perturbations" introduced via $[K]$.

We proceed now with the formal development. A <u>generalized saddle program</u> $S(K)$ <u>on</u> $R^m \times R^n$ <u>with perturbations in</u> $R^p \times R^q$ is a closed proper saddle-function K on $(R^p \times R^m) \times (R^q \times R^n)$. Each saddle-function $\tilde{K}(o,\cdot,o,\cdot)$ on $R^m \times R^n$ for $\tilde{K}$ in $[K]$ is called an <u>objective function</u> of $S(K)$. The particular functions $\underline{K}(o,\cdot,o,\cdot)$ and $\overline{K}(o,\cdot,o,\cdot)$ are called the <u>lower</u> and <u>upper</u> objective functions, respectively. (Recall the convention established in Theorem 0.1: for a closed saddle-function K, its convex-closure is denoted by $\underline{K}$ and its concave-closure by $\overline{K}$.) A pair

(x,y) is a <u>feasible solution</u> of $S(K)$ iff it is in the domain of every objective function of $S(K)$. It is not hard to show that this is equivalent to the condition that $(o,x,o,y) \in \operatorname{dom} K$. The <u>optimal value in</u> $S(K)$ <u>exists</u> (<u>and equals</u> α) iff the saddle-values of all the objective functions of $S(K)$ exist finitely and are equal (to α). A pair (x,y) is an <u>optimal solution</u> of $S(K)$ iff (x,y) is a saddle-point of every objective function of $S(K)$ and $\underline{K}(o,x,o,y) = \overline{K}(o,x,o,y) \in R$. It is not hard to show that if (x,y) is an optimal solution, then it is a feasible solution and the optimal value exists and equals $\tilde{K}(o,x,o,y)$ for any $\tilde{K}$ in $[K]$.

The program $S(K)$ is consistent (respectively <u>strongly consistent</u>) iff there exist points x and y such that $(o,x,o,y) \in \operatorname{dom} K$ (respectively $(o,x,o,y) \in \operatorname{ri}(\operatorname{dom} K)$). Thus, $S(K)$ is consistent iff it has a feasible solution. Also, $S(K)$ is consistent whenever the optimal value in $S(K)$ exists.

We say that $S(K)$ <u>has a well-defined primal problem</u> iff all the objective functions belong to the same equivalence class, which we denote by $[K_0]$. In this event the definitions of feasible solutions, optimal value and optimal solutions of $S(K)$ can be simplified, since equivalent saddle-functions have the same domain, saddle-value and saddle-points. By Theorem 6.2 below, if $S(K)$ is strongly consistent then it has a well-defined primal problem which is in fact given by a closed proper equivalence class. More generally, for any (u,v) we say that <u>the perturbation</u> (u,v) <u>in</u> $S(K)$ <u>is well-defined</u> iff the saddle-functions $\tilde{K}(u,\cdot,v,\cdot)$ on $R^m \times R^n$, for $\tilde{K}$ in $[K]$, all belong to a single equivalence class, which we denote by $[K_{u,v}]$. Thus, $S(K)$ has a well-defined primal problem iff the perturbation (o,o) in $S(K)$ is well-defined (in which case $[K_{o,o}]$ is denoted simply by $[K_0]$).

Suppose $S(K)$ is a generalized saddle program on $R^m \times R^n$ with

perturbations in $R^p \times R^q$, and let $[L]$ be the equivalence class of closed proper saddle-functions obtained from $[K]$ via the relation

$$L(s,z,t,w) = -K^*(-z,s,-w,t).$$

The generalized saddle program $S(L)$ on $R^p \times R^q$ with perturbations in $R^m \times R^n$ is the dual of $S(K)$. It is easy to show that the dual of $S(L)$ is $S(K)$. The program $S(K)$ has a well-defined dual problem iff the dual program $S(L)$ has a well-defined primal problem $[L_0]$, and in this event the dual problem of $S(K)$ is the minimax problem given by $[L_0]$. Example 6.3 shows that a generalized saddle program can even be strongly consistent without having a well-defined dual problem. However, Lemma 6.4 furnishes conditions on $S(K)$ which ensure that the dual problem is well-defined.

For the remainder of §6 let $S(K)$ and $S(L)$ be dual programs, where for definiteness K is assumed to be concave-convex on $(R^p \times R^m) \times (R^q \times R^n)$. Thus, L is convex-concave on $(R^m \times R^p) \times (R^n \times R^q)$. Also, let concave-convex functions P_1 and P_2 be defined on $R^p \times R^q$ by

$$P_1(u,v) = \sup_x \inf_y \underline{K}(u,x,v,y)$$

and

$$P_2(u,v) = \inf_y \sup_x \overline{K}(u,x,v,y),$$

and let convex-concave functions Q_1 and Q_2 be defined on $R^m \times R^n$ by

$$Q_1(s,t) = \sup_w \inf_z \underline{L}(s,z,t,w)$$

and

$$Q_2(s,t) = \inf_z \sup_w \overline{L}(s,z,t,w).$$

Finally, let linear transformations $A_1: R^p \times R^m \to R^p$, $A_2: R^q \times R^n \to R^q$, $B_1: R^m \times R^p \to R^m$ and $B_2: R^n \times R^q \to R^n$ be defined by

$$A_1(u,x) = u, \quad B_1(s,z) = s,$$

$$A_2(v,y) = v, \quad B_2(t,w) = t,$$

and put $A = A_1 \times A_2$ and $B = B_1 \times B_2$. Observe that $A^* = A_1^* \times A_2^*$ and

$B^* = B_1^* \times B_2^*$, where

$$A_1^*(z) = (z,o), \quad B_1^*(x) = (x,o)$$

$$A_2^*(w) = (w,o), \quad B_2^*(y) = (y,o).$$

The saddle-functions P_1 and P_2 are called the <u>lower</u> and <u>upper perturbation functions</u> of $S(K)$, respectively. A pair (z,w) is a <u>Kuhn–Tucker vector</u> for $S(K)$ iff

$$P_1(o,o) = P_2(o,o) = \alpha \in R$$

and

$$\langle u,z\rangle + P_2(u,o) \leq \alpha \leq P_1(o,v) + \langle v,w\rangle$$

for each (u,v). Observe that $P_1(o,o) = P_2(o,o) = \alpha \in R$ iff the optimal value in $S(K)$ exists and equals α. It is not hard to show using (37.4.1) that if P_1 and P_2 belong to a proper equivalence class $[P]$, then (z,w) is a Kuhn–Tucker vector for $S(K)$ iff $-(z,w) \in \partial P(o,o)$. Kuhn–Tucker vectors for $S(L)$ are defined similarly by using the lower and upper perturbation functions of $S(L)$, i.e., Q_1 and Q_2. These Kuhn–Tucker vectors can be interpreted as generalized "equilibrium price vectors" in much the same way as in [44, pp. 295-296].

The following example shows that the foregoing framework of dual pairs of generalized saddle-programs includes as a special case Rockafellar's dual pairs of generalized convex programs.

EXAMPLE 6.1. Let $F\colon R^q \to R^n$ be a closed proper convex bifunction, and let (P) and (P*) denote the generalized convex program and its dual which correspond to F and its adjoint bifunction $F^*\colon R^n \to R^q$. Define $K(u,x,v,y) = (Fv)(y)$ for every $(u,x) \in R^p \times R^m$ and $(v,y) \in R^q \times R^n$ (here p and m can be any positive integers). Then K is a closed proper concave-convex function. It can be verified as an instructive exercise that the concepts defined above for the program $S(K)$ and its dual $S(L)$ exactly

"mirror" the like-named concepts from Rockafellar [44] for (P) and (P*). For example, S(K) (resp. S(L)) is consistent or strongly consistent according as (P) (resp. (P*)) is consistent or strongly consistent; and so on. Furthermore, it can be seen that the Lagrangian saddle-function M associated with S(K) and S(L), which will be introduced following Theorem 6.8, exactly mirrors the Lagrangian associated with (P) and (P*). The fact that all the various concepts associated with (P) and (P*) are reflected in this program S(K) and its dual furnishes a general means of converting examples from convex programming into examples in saddle programming which exhibit the corresponding pathological behavior.

THEOREM 6.2. *Assume* $(u,v) \in A\ \mathrm{ri}(\mathrm{dom}\ K)$, i.e. *assume there exist points* x *and* y *such that* $(u,x,v,y) \in \mathrm{ri}(\mathrm{dom}\ K)$. *Then the perturbation* (u,v) *in* S(K) *is well-defined. In fact, the equivalence class* $[K_{u,v}]$ *is closed and proper with least and greatest members* $\underline{K}(u,\cdot,v,\cdot)$ *and* $\overline{K}(u,\cdot,v,\cdot)$ *respectively, and*

$$\mathrm{ri}(\mathrm{dom}\ K_{u,v}) = \{(x,y) \mid (u,x,v,y) \in \mathrm{ri}(\mathrm{dom}\ K)\}$$

where "ri" *can be deleted or replaced by* "cl" *throughout the identity.*

PROOF. Define linear transformations $T_1 \colon R^m \to R^p \times R^m$ and $T_2 \colon R^n \to R^q \times R^n$ by $T_1(x) = (o,x)$ and $T_2(y) = (o,y)$, and put $T = T_1 \times T_2$. Define a function H by

$$H(u',x',v',y') = K(u'+u,x',v'+v,y').$$

Clearly, H is closed proper concave-convex and $\mathrm{dom}\ H = \mathrm{dom}\ K - \{(u,o,v,o)\}$. Thus the hypothesis on (u,v) is equivalent to $\mathrm{range}\ T \cap \mathrm{ri}(\mathrm{dom}\ H) \neq \emptyset$, and hence Theorem 2.2 implies various facts about the equivalence class $[HT]$. Since $HT = K(u,\cdot,v,\cdot)$, these facts convert easily into the assertions about $[K_{u,v}]$. The formulas for $\mathrm{ri}(\mathrm{dom}\ K_{u,v})$ and $\mathrm{cl}(\mathrm{dom}\ K_{u,v})$ follow by (6.7).

COROLLARY 6.2.1. *Assume* S(K) *is strongly consistent and that*

rec $\text{cone}_j K_0$ <u>is a subspace for</u> $j = 1$ <u>and</u> 2. <u>Then there exists an optimal solution of</u> $S(K)$.

PROOF. By Theorem 6.2, $S(K)$ has a well-defined primal problem and $[K_0]$ is closed and proper. By Lemma 0.4, $(o,o) \in \text{ri}(\text{dom}(K_0)^*)$. Hence (37.5.3) implies K_0 has a saddle-point.

Before proceeding any further, it might be well to illustrate some of the pathology which is possible in a dual pair of generalized saddle programs. The next example exhibits a program $S(K)$ having the following properties: (1) every perturbation in $S(K)$ is well-defined (so <u>a fortiori</u> $S(K)$ has a well-defined primal problem); (2) the lower and upper perturbation functions of $S(K)$ fail to be equivalent; (3) the dual program is consistent; and (4) $S(K)$ fails to have a well-defined dual problem.

EXAMPLE 6.3. Suppose J is a closed proper concave-convex function on $R^m \times R^n$. Put $p = m$ and $q = n$ and define $K(u,x,v,y) = J(x - u, y - v)$. Let T_1 and T_2 be linear transformations given by $T_1(u,x) = x - u$ and $T_2(v,y) = y - v$, and put $T = T_1 \times T_2$. Since $\text{range } T \cap \text{ri}(\text{dom } J) \neq \emptyset$ trivially, Theorem 2.2 implies that $K = JT$ is closed and proper with $\text{ri}(\text{dom } K) = T^{-1}\text{ri}(\text{dom } J)$. By Theorem 6.2 it follows that for each $(u,v) \in R^p \times R^q$ the perturbation (u,v) in the program $S(K)$ is well-defined. It is easy to compute that $P_1(u,v) = \sup \inf \underline{J} = -\overline{J}^*(o,o)$ and $P_2(u,v) = \inf \sup \overline{J} = -\underline{J}^*(o,o)$. Hence $P_1 \sim P_2$ iff $\underline{J}^*(o,o) = \overline{J}^*(o,o)$. Now suppose J is such that $\text{dom } J^*$ is bounded. Then Lemma 3.6 and Theorem 3.5 imply that $[K^*] = [T^*J^*]$, $\text{dom } K^* = T^*\text{dom } J^*$, and (since $T_1^*(s) = (-s,s)$, $T_2^*(t) = (-t,t)$) the least and greatest members of $[K^*]$ are

$$\underline{K}^*(z,s,w,t) = \sup_{\{s|-s=z\}} \inf_{\{t|-t=w\}} \underline{J}^*(x,t)$$

and

$$\overline{K}^*(z,s,w,t) = \inf_{\{t|-t=w\}} \sup_{\{s|-s=z\}} \overline{J}^*(s,t).$$

Since $\underline{L}(s,z,t,w) = -\overline{K}^*(-z,s,-w,t)$ and $\overline{L}(s,z,t,w) = -\underline{K}^*(-z,s,-w,t)$, this implies that

$$\underline{L}(s,z,t,w) = \begin{cases} -\overline{J}^*(s,t) & \text{if } s = z \text{ and } t = w \\ +\infty & \text{if } s \neq z \text{ and } t = w \\ -\infty & \text{if } t \neq w \end{cases}$$

and

$$\overline{L}(s,z,t,w) = \begin{cases} -\underline{J}^*(s,t) & \text{if } s = z \text{ and } t = w \\ -\infty & \text{if } s = z \text{ and } t \neq w \\ +\infty & \text{if } s \neq z \end{cases}$$

From these formulas it follows that, for each $(s,t) \in \text{dom } J^*$, the perturbation (s,t) in $S(L)$ is well-defined iff $\underline{J}^*(s,t) = \overline{J}^*(s,t)$. In view of all these facts, in order to obtain properties (1) through (4) we need only specify a J such that $\text{dom } J^*$ is bounded, $(o,o) \in \text{dom } J^*$, and $\underline{J}^*(o,o) \neq \overline{J}^*(o,o)$. It suffices to take $[J]$ to be the conjugate of the equivalence class used in Examples 2.10 and 2.11.

While we are concerned mainly with applying the results of §§2, 3 and 5 to dual pairs of generalized saddle programs, we note here some of the results that follow from §1. From Lemma 1.7 it can be deduced that in general

$$\underline{L}(o,\cdot,o,\cdot) \leq (\underline{-P_1})^* \leq (\overline{-P_2})^* \leq \overline{L}(o,\cdot,o,\cdot)$$

and dually

$$\underline{K}(o,\cdot,o,\cdot) \leq (\underline{-Q_1})^* \leq (\overline{-Q_2})^* \leq \overline{K}(o,\cdot,o,\cdot).$$

If $S(L)$ has a well-defined primal problem given by $[L_0]$ and $[L_0]$ is closed, then it can be deduced from Theorem 1.1 that each concave-convex function P satisfying $P_1 \leq P \leq P_2$ is simple and satisfies

$$cl_2 cl_1 P = -(\overline{L_0})^*, \quad cl_1 cl_2 P = -(\underline{L_0})^*,$$
$$\text{dom}(cl_2 cl_1 P) = \text{dom}(L_0)^* = \text{dom}(cl_1 cl_2 P),$$

where

$$\text{dom}(L_0)^* \subset \text{cl}\{(u,v) \mid (u,x,v,y) \in \text{dom } K \text{ for some } x,y\}.$$

If in addition $[L_0]$ is proper (i.e. $S(L)$ is consistent), then P is also proper and has the same kernel as $-(L_0)^*$.

By Theorem 6.2, $S(K)$ has a well-defined dual problem $[L_0]$ whenever $S(L)$ is strongly consistent. The next lemma dualizes this useful condition.

LEMMA 6.4. $S(L)$ *is strongly consistent iff*

$$(\text{rec}_1 K)(o,x) \geq 0 \quad \textit{implies} \quad (\text{rec}_1 K)(o,-x) \geq 0$$

and

$$(\text{rec}_2 K)(o,y) \leq 0 \quad \textit{implies} \quad (\text{rec}_2 K)(o,-y) \leq 0.$$

PROOF. Observe that $\underline{L}(o,z,o,w) = -\overline{K}^*A^*(-z,-w)$. Hence $S(L)$ is strongly consistent iff range $A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$. Now apply the equivalence between (a) and (c) of Lemma 2.9.

THEOREM 6.5. *Assume* $S(L)$ *is strongly consistent. Then* P_1 *and* P_2 *belong to the closed proper equivalence class* $[P] = [-(L_0)^*]$ *and* $\text{dom } P \subset A \text{ dom } K$.

PROOF. By Theorem 6.2, $\underline{L}(o,\cdot,o,\cdot)$ is the least member of $[L_0]$, which is closed and proper. Hence $-\underline{L}(o,-z,o,-w) = -\underline{L_0}(-z,-w) = (\underline{-(L_0)^*})^*(z,w)$. But as noted in the proof of Lemma 6.4, $S(L)$ is strongly consistent iff range $A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$, and $\overline{K}^*A^*(z,w) = -\underline{L}(o,-z,o,-w)$. Hence Theorem 2.8 implies that the equivalence class $[AK]$ is well-defined and equals $[-(L_0)^*]$, and $\text{dom } AK \subset A \text{ dom } K$. Now observe that

$$P_1(u,v) = \sup_{A_1^{-1}(u)} \inf_{A_2^{-1}(v)} \underline{K}, \quad P_2(u,v) = \inf_{A_2^{-1}(v)} \sup_{A_1^{-1}(u)} \overline{K}.$$

Thus P_1 and P_2 belong to $[AK]$. Taking $[P] = [AK]$, the theorem follows.

COROLLARY 6.5.1. *Assume* $S(L)$ *is strongly consistent. Then the following conditions on* $(z,w) \in R^p \times R^q$ *are equivalent:*

(i) (z,w) *is an optimal solution of* $S(L)$;

(ii) (z,w) *is a Kuhn-Tucker vector for* $S(K)$;

(iii) $-(z,w) \in \partial P(o,o)$;

(iv) $(-z,o,-w,o) \in \partial K(o,x,o,y)$ *for some* $(x,y) \in R^m \times R^n$.

PROOF. By (37.5) and Theorem 6.5, (i) is equivalent to $(z,w) \in \partial(-P)(o,o)$, which is equivalent to (iii). Since P_1 and P_2 belong to $[P]$, (37.4.1) implies that $\partial P(o,o) = \partial_1 P_2(o,o) \times \partial_2 P_1(o,o)$ and $P_1(o,o) = P_2(o,o) = \alpha$. Also, (37.4) implies $\operatorname{dom} \partial P \subset \operatorname{dom} P$, so that α is finite. From these facts it follows that (iii) is equivalent to (ii). Finally, observe that (37.5) implies (iii) is equivalent to $(o,o) \in \partial P^*(-z,-w)$. Since $[P^*] = [K^*A^*]$ by the proof of Theorem 6.5, Theorem 2.3 implies that $\partial P^*(-z,-w) = A\partial K^*(A^*(-z,-w))$. Hence $(o,o) \in \partial P^*(-z,-w)$ is equivalent to the existence of $(u,x,v,y) \in \partial K^*(-z,o,-w,o)$ such that $A_1(u,x) = 0$ and $A_2(v,y) = 0$. By the definitions of A_1 and A_2 and (37.5), this last condition is equivalent to (iv).

COROLLARY 6.5.2. *Assume* $S(L)$ *is strongly consistent, and let* $[P]$ *be the equivalence class containing* P_1 *and* P_2. *Then*

$$\sup\inf L_0 = \underline{P}(o,o) \le \overline{P}(o,o) = \inf\sup L_0 .$$

Furthermore, for any $\tilde{P} \in [P]$,

$$\sup\inf L_0 = \liminf_{v \to 0} \tilde{P}(o,v)$$

whenever $-\infty < \sup\inf L_0$ *or* $\tilde{P}(o,o) < +\infty$, *and*

$$\limsup_{u \to 0} \tilde{P}(u,o) = \inf\sup L_0$$

whenever $-\infty < \tilde{P}(o,o)$ *or* $\inf\sup L_0 < +\infty$.

PROOF. Clearly $\sup\inf L_0 = -(\overline{L_0})^*(o,o)$ which equals $\underline{P}(o,o)$ by Theorem 6.5. Similarly, $\inf\sup L_0 = -(\underline{L_0})^*(o,o) = \overline{P}(o,o)$. Now let $\tilde{P} \in [P]$ be given. By Theorem 0.1, $\underline{P}(o,o) = (\mathrm{cl}_2\tilde{P})(o,o)$, which by definition equals $(\mathrm{cl}\,\tilde{P}(o,\cdot))(o)$. Now in general, for a convex function f one has $(\mathrm{cl}\,f)(x) = \liminf_{y \to x} f(y)$ except when $(\mathrm{cl}\,f)(x) = -\infty$ and $f(x) = +\infty$. Applying this fact

to the case at hand yields that $\sup\inf L_0 = (\mathrm{cl}\ \tilde{P}(o,\cdot))(o) = \liminf_{v \to o} \tilde{P}(o,v)$ except when $\sup\inf L_0 = -\infty$ and $\tilde{P}(o,o) = +\infty$. The remaining assertion follows similarly.

COROLLARY 6.5.3. Assume $S(L)$ is strongly consistent. If the optimal value in $S(L)$ exists and equals α, then the optimal value in $S(K)$ exists and equals α.

PROOF. Since the saddle-functions $\tilde{L}(o,\cdot,o,\cdot)$ for $\tilde{L}$ in $[L]$ are all equivalent to L_0, $\sup\inf L_0 = Q_1(o,o)$ and $\inf\sup L_0 = Q_2(o,o)$. Also, the optimal value in $S(L)$ exists and equals α iff $Q_1(o,o) = Q_2(o,o) = \alpha \in R$. Since $\underline{P} \le P_1 \le P_2 \le \overline{P}$, the assertion now follows from Corollary 6.5.2.

COROLLARY 6.5.4. Assume $S(K)$ is strongly consistent. In order that

$$\mathrm{ri}(B\ \mathrm{dom}\ L) \subset \mathrm{dom}\ K_0^* \subset B\ \mathrm{dom}\ L,$$

it is necessary and sufficient that

$$(\mathrm{rec}_1 K)(o,x) = (\mathrm{rec}_1 K_0)(x),\ \forall x \in R^m$$

and

$$(\mathrm{rec}_2 K)(o,y) = (\mathrm{rec}_2 K_0)(y),\ \forall y \in R^n.$$

PROOF. Dualizing the proof of Theorem 6.5 shows that range $B^* \cap \mathrm{ri}(\mathrm{dom}\ L^*) \ne \emptyset$ and $\mathrm{dom}\ K_0^* = \mathrm{dom}\ B^! \subset B\ \mathrm{dom}\ L$. Hence by Theorem 2.6 and (6.3.1), $\mathrm{ri}(B\ \mathrm{dom}\ L) \subset \mathrm{dom}\ K_0^* \subset B\ \mathrm{dom}\ L$ occurs iff $\mathrm{rec}_j(L^*B^*) = (\mathrm{rec}_j L^*)B_j^*$ for $j = 1$ and 2. But the identities $-(\mathrm{rec}_1 K)(o,-x) = (\mathrm{rec}_1 L^*)(B_1^* x)$ and $-(\mathrm{rec}_1 K_0)(-x) = (\mathrm{rec}_1 L^*B^*)(x)$ can be verified, along with similar identities for $j = 2$. The corollary then follows.

Example 6.14 demonstrates that the domain inclusion in Theorem 6.5 cannot be strengthened to equality without a stronger hypothesis. For any product set $C \times D$ contained in dom K, let $\mathrm{Hyp}(C \times D)$ denote the following hypothesis:

$$\forall (u,x) \in C,\ (\mathrm{rec}\ \underline{K}(u,x,\cdot,\cdot))(o,y) \le 0 \Rightarrow (\mathrm{rec}\ \underline{K}(u,x,\cdot,\cdot))(o,-y) \le 0;$$

and

$$\forall (v,y) \in D,\ (\operatorname{rec} \overline{K}(\cdot,\cdot,v,y))(o,x) \geq 0 \Rightarrow (\operatorname{rec} \overline{K}(\cdot,\cdot,v,y))(o,-x) \geq 0.$$

If $\mathrm{Hyp}(C \times D)$ for some $C \times D \supset \mathrm{ri}(\mathrm{dom}\, K)$, then Lemma 6.4 implies that $S(L)$ is strongly consistent.

LEMMA 6.6. *The following conditions are equivalent, and they imply* $\mathrm{Hyp}(\mathrm{dom}\, K)$:

$$\text{(a)} \quad \begin{cases} (o,x) \in o^+(\mathrm{cl}(\mathrm{dom}_1 K)) \Rightarrow x = o \\ (o,y) \in o^+(\mathrm{cl}(\mathrm{dom}_2 K)) \Rightarrow y = o \end{cases}$$

(b) *There exist points* u *and* v *such that the sets* $(\{u\} \times R^m) \cap \mathrm{ri}(\mathrm{dom}_1 K)$ *and* $(\{v\} \times R^n) \cap \mathrm{ri}(\mathrm{dom}_2 K)$ *are nonempty and bounded.*

PROOF. By Lemma 3.6.

THEOREM 6.7. *Assume* $\mathrm{Hyp}(\mathrm{ri}(\mathrm{dom}\, K))$. *Then the conclusions of Theorem 6.5 hold, and in addition* $A\, \mathrm{ri}(\mathrm{dom}\, K) \subset \mathrm{dom}\, P \subset A\, \mathrm{dom}\, K$. *For each* $(u,v) \in A\, \mathrm{ri}(\mathrm{dom}\, K)$, *the perturbation* (u,v) *in* $S(K)$ *is well-defined,* $[K_{u,v}]$ *is closed and proper, and the set of saddle-points of* $[K_{u,v}]$ *is nonempty. Each such saddle-point* (x,y) *satisfies* $(u,x,v,y) \in \mathrm{dom}\, \partial K$ *and* $\tilde{P}(u,v) = \tilde{K}(u,x,v,y)$ *for every* $\tilde{P} \in [P]$ *and* $\tilde{K} \in [K]$.

PROOF. By Lemma 6.4, $\mathrm{Hyp}(\mathrm{ri}(\mathrm{dom}\, K))$ implies $S(L)$ is strongly consistent. Hence the conclusions of Theorem 6.5 hold, and in particular $\mathrm{dom}_1 P_1 \times \mathrm{dom}_2 P_2 = \mathrm{dom}\, P$. Thus it follows from $\mathrm{Hyp}(\mathrm{ri}(\mathrm{dom}\, K))$ and Lemma 3.3 that $A\, \mathrm{ri}(\mathrm{dom}\, K) \subset \mathrm{dom}\, P$. If $(u,v) \in A\, \mathrm{ri}(\mathrm{dom}\, K)$, then Theorem 6.2 implies the perturbation (u,v) is well-defined and that $[K_{u,v}]$ is closed and proper. The assertions concerning the saddle-points of $[K_{u,v}]$ are immediate from $\mathrm{Hyp}(\mathrm{ri}(\mathrm{dom}\, K))$ and Theorem 3.4.

If actually $\mathrm{Hyp}(\mathrm{dom}\, K)$, then Lemma 3.3 and Theorem 3.5 imply that $\mathrm{dom}\, P = A\, \mathrm{dom}\, K$ and that P_1 and P_2 are the least and greatest members of $[P]$, respectively.

COROLLARY 6.7.1. _Assume $S(K)$ is strongly consistent and either $\operatorname{dom} K_0$ is bounded or $\operatorname{Hyp}(\operatorname{ri}(\operatorname{dom} K))$. Then there exist optimal solutions of both $S(K)$ and $S(L)$, and the optimal values in the two programs are equal. Moreover, the optimal solutions of one program are precisely the Kuhn-Tucker vectors for the other._

PROOF. By Theorem 6.2 and Lemma 6.6, if $S(K)$ is strongly consistent and $\operatorname{dom} K_0$ is bounded, then $\operatorname{Hyp}(\operatorname{dom} K)$, and hence $\operatorname{Hyp}(\operatorname{ri}(\operatorname{dom} K))$. By Theorem 6.7 and (6.3.1), $S(K)$ is strongly consistent iff $(c,o) \in \operatorname{ri}(\operatorname{dom} P) = A \operatorname{ri}(\operatorname{dom} K)$. Hence Theorem 6.7 implies there exists a saddle-point of $[K_0]$, i.e. an optimal solution of $S(K)$. Also, $\operatorname{ri}(\operatorname{dom} P) \subset \operatorname{dom} \partial P$ by (37.4), so that $\partial P(o,c) \neq \emptyset$. By Corollary 6.5.1 this implies there exists an optimal solution of $S(L)$. Since both programs are strongly consistent, Corollary 6.5.1 implies that the optimal solutions of one are the Kuhn-Tucker vectors for the other. The two optimal values are equal by Corollary 6.5.3.

As a criterion guaranteeing the existence of an optimal solution of $S(K)$, the hypothesis of Corollary 6.7.1 is stronger than necessary. This is clear from the next lemma and Corollary 6.2.1. In fact, as will become clear later in the section, the hypothesis of Corollary 6.2.1 suffices for the first assertion of Corollary 6.7.1.

LEMMA 6.8. _Assume $S(K)$ is strongly consistent. Then $\operatorname{Hyp}(\operatorname{ri}(\operatorname{dom} K))$ implies that $\operatorname{rec\,cone}_j K_0$ is a subspace for $j = 1$ and 2._

PROOF. Let T be as in the proof of Theorem 6.2. Then strong consistency of $S(K)$ is equivalent to $\operatorname{range} T \cap \operatorname{ri}(\operatorname{dom} K) \neq \emptyset$, and hence $[K_0] = [KT]$ by Theorem 2.2. Therefore the formulas in Lemma 2.5 can be applied to show that

$$\operatorname{rec\,cone}_1 K_0 = \{x \mid (o,x) \in \operatorname{rec\,cone} \bar{K}(\cdot,\cdot,o,y) \text{ whenever } (o,y) \in \operatorname{ri}(\operatorname{dom}_2 K)\}$$

and

$\text{rec cone}_2 K_0 = \{y \mid (o,y) \in \text{rec cone } \underline{K}(o,x,\cdot,\cdot) \text{ whenever } (o,x) \in ri(\text{dom}_1 K)\}$. For these two sets to be subspaces it suffices to show they are closed under scaler multiplication by -1, and this follows immediately from strong consistency and Hyp(ri(dom K)).

Let M be defined on $(R^m \times R^q) \times (R^n \times R^p)$ by

$$M(x,w,y,z) = \sup_u \inf_v \{\langle u,z\rangle + \langle v,w\rangle + K(u,x,v,y)\}.$$

Then $M(x,w,y,z) = -W(-z,y,-w,x)$, where W belongs to the partial conjugate of $[K]$ in u and v. Hence it follows from Theorem 5.1 that M is closed proper concave-convex and depends only on $[K]$, and that M is polyhedral whenever K is polyhedral. The equivalence class containing M is called the <u>Lagrangian</u> of $S(K)$. Similarly, the Lagrangian of $S(L)$ is the equivalence class containing the function N given by

$$N(x,w,y,z) = \sup_t \inf_s \{\langle s,x\rangle + \langle t,y\rangle + L(s,z,t,w)\}.$$

In view of the next theorem, $[M]$ is called <u>the</u> Lagrangian of the dual pair $S(K)$, $S(L)$. From the fact that the partial conjugacy operation induces a symmetric one-to-one correspondence among closed proper equivalence classes, it follows that a dual pair of generalized saddle programs is completely determined by its Lagrangian.

THEOREM 6.9. <u>The saddle-functions</u> M <u>and</u> N <u>are equivalent</u>.

PROOF. Since $[L]$ is obtained from $[K^*]$ via the relation $L(s,z,t,w) = -K^*(-z,s,-w,t)$, it follows by (36.1) and Theorem 0.1(b) that $[L]$ contains the function L given by

$$L(s,z,t,w) = -\sup_y \inf_x \inf_u \sup_v \{\langle u,-z\rangle + \langle x,s\rangle + \langle v,-w\rangle + \langle y,t\rangle - K(u,x,v,y)\}$$

$$= -\sup_y \inf_x \{\langle x,s\rangle + \langle y,t\rangle - M(x,w,y,z)\}.$$

If $[H]$ denotes the partial conjugate of $[M]$ in x and y, this means $[L] = [-H]$. Now by Theorem 5.1 the function N depends only on $[L]$. Hence

N is equivalent to the function $\tilde{N}$ given by

$$\tilde{N}(x,w,y,z) = \sup_t \inf_s \{\langle s,x\rangle + \langle t,y\rangle - H(s,z,t,w)\}.$$

But $\tilde{N}$ belongs to the partial conjugate of $[H]$ in s and t, and by Theorem 5.2 this is the same as $[M]$. This shows that $\tilde{N} \in [M]$, and hence $[N] = [M]$.

The next theorem says that, up to a reordering of the variables, the conjugate of the Lagrangian of $S(K)$ coincides with both the partial conjugate of $[K]$ in x and y and the partial conjugate of $[L]$ in z and w.

THEOREM 6.10. <u>Let $[M]$ be the Lagrangian of $S(K)$, let H be any member of the partial conjugate of $[K]$ in x and y, and let J be any member of the partial conjugate of $[L]$ in z and w. Then $[M^*]$ contains the functions</u>

$$(s,v,t,u) \to H(v,s,u,t)$$

and

$$(s,v,t,u) \to J(t,u,s,v).$$

PROOF Recall that $[M]$ contains $(x,w,y,z) \to -W(-z,y,-w,x)$, where W is a member of the partial conjugate of $[K]$ in u and v. Hence (using (36.1) and Theorem 0.1(b) to interchange "sup" with "inf"), $[M^*]$ contains

$$(s,v,t,u) \to \sup_y \inf_x \sup_z \inf_w \{\langle x,s\rangle + \langle w,v\rangle + \langle y,t\rangle + \langle z,u\rangle + W(-z,y,-w,x)\}.$$

Now observe that $(u,x,v,y) \to \inf_z \sup_w \{\langle -z,u\rangle + \langle -w,v\rangle - W(-z,y,-w,x)\}$ is a member of the partial conjugate of $[W]$ in z and w, which by Theorem 5.2 is just $[K]$. Hence M^* contains $(s,v,t,u) \to$ $\sup_y \inf_x \{\langle x,s\rangle + \langle y,t\rangle \; \tilde{K}(u,x,v,y)\}$ for some $\tilde{K} \in [K]$. But up to reordering the variables this function belongs to the partial conjugate of $[K]$ in x and y. This shows that $[M^*]$ contains $(s,v,t,u) \to H(v,s,u,t)$. Similarly, $[M^*]$ contains

$$(s,v,t,u) \to \sup_z \inf_w \sup_y \inf_x \{\langle x,s\rangle + \langle w,v\rangle + \langle y,t\rangle + \langle z,u\rangle - H(x,w,y,z)\},$$

where $M(x,w,y,z) = \sup_{\bar{u}} \inf_{\bar{v}} \{\langle \bar{u},z\rangle + \langle \bar{v},w\rangle + K(\bar{u},x,\bar{v},y)\}$. Since

$$(z,s,w,t) \to \sup_{y} \inf_{x} \inf_{\bar{u}} \sup_{\bar{v}} \{\langle \bar{u},z\rangle + \langle x,s\rangle + \langle \bar{v},w\rangle + \langle y,t\rangle - K(\bar{u},x,\bar{v},y)\}$$

belongs to $[K^*]$, this means that $[M^*]$ contains

$$(s,v,t,u) \to \sup_{z} \inf_{w} \{\langle z,u\rangle + \langle w,v\rangle + \tilde{K}^*(-z,s,-w,t)\}$$

for some $\tilde{K}^* \in [K^*]$, i.e.

$$(s,v,t,u) \to \sup_{z} \inf_{w} \{\langle z,u\rangle + \langle w,v\rangle - \tilde{L}(s,z,t,w)\}$$

for some $\tilde{L} \in [L]$. But up to reordering the variables this function belongs to the partial conjugate of $[L]$ in z and w. This shows that $[M^*]$ contains $(s,v,t,u) \to J(t,u,s,v)$.

COROLLARY 6.10.1. *If the saddle-value of the Lagrangian exists and equals* α, *where* $\alpha \in R$, *then the optimal values in* $S(K)$ *and* $S(L)$ *exist and equal* α.

PROOF. The saddle-value of M exists and equals α iff $\underline{M}^*(o,o,o,o) = \bar{M}^*(o,o,o,o) = -\alpha$. By Theorem 0.1(b) this is equivalent to $\tilde{M}^*(o,o,o,o) = -\alpha$ for every $\tilde{M}^* \in [M^*]$. For $i = 1$ and 2, $-P_i(u,v) = H_i(v,o,u,o)$ for a certain member H_i of the partial conjugate of $[K]$ in x and y, and $-Q_i(s,t) = J_i(t,o,s,o)$ for a certain member J_i of the partial conjugate of $[L]$ in z and w. Hence Theorem 6.10 implies that $-P_i(o,o) = -\alpha$ $-Q_i(o,o) = -\alpha$ for $i = 1,2$.

COROLLARY 6.10.2. *If there exist points* u *and* v *such that* $(o,v,o,u) \in \mathrm{ri}(\mathrm{dom}\, M^*)$, *then* $-P_1$ *and* $-P_2$ *belong to a closed proper equivalence class which contains the upper and lower conjugate of every objective function of* $S(L)$. *Dually, if there exist points* s *and* t *such that* $(s,o,t,o) \in \mathrm{ri}(\mathrm{dom}\, M^*)$, *then* $-Q_1$ *and* $-Q_2$ *belong to a closed proper equivalence class which contains the upper and lower conjugate of every objective function of* $S(K)$.

PROOF. Assume $(o,v,o,u) \in \text{ri}(\text{dom } M^*)$ for some u and v. Then Theorem 6.2 implies that the functions $(v,u) \to \tilde{M}^*(o,v,o,u)$ for $\tilde{M}^* \in [M^*]$ all belong to a single closed proper equivalence class. By Theorem 6.10 this implies that the functions

$$(v,u) \to -P_1(u,v) = \inf_x \sup_y \{\langle x,o\rangle + \langle y,o\rangle - \underline{K}(u,x,v,y)\},$$

$$(v,u) \to -P_2(u,v) = \sup_y \inf_x \{\langle x,o\rangle + \langle y,o\rangle - \bar{K}(u,x,v,y)\},$$

$$(v,u) \to (\overline{\bar{L}(o,\cdot,o,\cdot)})^*(u,v) = \inf_w \sup_z \{\langle z,u\rangle + \langle w,v\rangle - \bar{L}(o,z,o,w)\},$$

$$(v,u) \to (\underline{\underline{L}(o,\cdot,o,\cdot)})^*(u,v) = \sup_z \inf_w \{\langle z,u\rangle + \langle w,v\rangle - \underline{L}(o,z,o,w)\}$$

are equivalent, closed and proper. The dual assertion follows similarly.

We call the conditions in the next lemma the <u>extremality conditions</u> associated with $S(K)$ and $S(L)$.

LEMMA 6.11. <u>For any</u> $(x,y) \in R^m \times R^n$ <u>and</u> $(z,w) \in R^p \times R^q$, <u>the following conditions are equivalent</u>:

(a) $(-z,o,-w,o) \in \partial K(o,x,o,y)$;

(b) $(-x,o,-y,o) \in \partial L(o,z,o,w)$;

(c) $(o,o,o,o) \in \partial M(x,w,y,z)$;

(d) (x,w,y,z) <u>is a saddle-point of the Lagrangian</u>.

PROOF. Observe that $-M(x,-w,y,-z) = W(z,y,w,x)$, where W is in the partial conjugate of $[K]$ in u and v. Also, Theorem 5.3 implies (a) is equivalent to $(o,o,o,o) \in \partial W(-z,y,-w,x)$. By (37.4) it follows that (a) is equivalent to (c). Trivially, (c) is equivalent to (d). Finally, (37.5) implies that (a) is equivalent to $(o,x,o,y) \in \partial K^*(-z,o,-w,o)$. By (37.4) and the relation $L(s,z,t,w) = -K^*(-z,s,-w,t)$ it follows that this last condition is equivalent to (b).

THEOREM 6.12. <u>Assume</u> $(x,y) \in R^m \times R^n$ <u>and</u> $(z,w) \in R^p \times R^q$ <u>satisfy the extremality conditions</u>. <u>Then</u> (x,y) <u>is an optimal solution of</u> $S(K)$ <u>and</u> (z,w) <u>is a Kuhn-Tucker vector for</u> $S(K)$. <u>Dually</u>, (z,w) <u>is an optimal</u>

<u>solution of</u> $S(L)$ <u>and</u> (x,y) <u>is a Kuhn–Tucker vector for</u> $S(L)$.

PROOF. By Lemma 6.11 we can suppose that $(-z,o,-w,o) \in \partial K(o,x,o,y)$. Then (37.4) and (37.4.1) imply that all the members of $[K]$ have the same finite value α at (o,x,o,y), and moreover

$$\tilde{K}(u,x',o,y) + \langle u,z\rangle \leq \alpha \leq \tilde{K}(o,x,v,y') + \langle v,w\rangle$$

for all $(u,x') \in R^p \times R^m$, $(v,y') \in R^q \times R^n$ and $\tilde{K} \in [K]$. Taking $u = o$ and $v = o$, this shows (x,y) is a saddle-point of $\tilde{K}(o,\cdot,o,\cdot)$ for each $\tilde{K} \in [K]$. Hence (x,y) is an optimal solution of $S(K)$, and $P_1(o,o) = \alpha = P_2(o,o)$. Also, by taking $\tilde{K} = \overline{K}$ we obtain $\langle u,z\rangle + \sup \overline{K}(u,\cdot,o,y) \leq \alpha$ for all $u \in R^p$, and by taking $\tilde{K} = \underline{K}$ we obtain $\alpha \leq \inf \underline{K}(o,x,v,\cdot) + \langle v,w\rangle$ for all $v \in R^q$. Hence (z,w) is a Kuhn–Tucker vector for $S(K)$. The dual assertion follows similarly, using the condition $(-x,o,-y,o) \in \partial L(o,z,o,w)$.

The next result is a generalization of the Kuhn–Tucker theorem. It is refined somewhat by Corollaries 6.17.2 and 6.17.3.

THEOREM 6.13. <u>Assume</u> $S(K)$ <u>is strongly consistent</u>. <u>Then the pair</u> $(x,y) \in R^m \times R^n$ <u>is an optimal solution of</u> $S(K)$ <u>iff there exists a pair</u> $(z,w) \in R^p \times R^q$ <u>such that</u> (x,w,y,z) <u>is a saddle-point of the Lagrangian</u>. <u>Such a pair</u> (z,w) <u>is a Kuhn–Tucker vector for</u> $S(K)$.

PROOF. By the dual version of Corollary 6.5.1, (x,y) is an optimal solution of $S(K)$ iff there exists a pair (z,w) such that $(-x,o,-y,o) \in \partial L(o,z,o,w)$. The theorem follows from this by Theorem 6.12.

By analogy with (36.6) for convex programming, one might ask whether the last assertion of Theorem 6.13 can be strengthened to the following: "Such pairs (z,w) are precisely the Kuhn–Tucker vectors for $S(K)$." The next example demonstrates that this does not hold in general. The reason is basically that the set $\partial M^*(o,o,o,o)$ of saddle-points of the Lagrangian is the product set

$$\partial M^*(\cdot,\cdot,o,o)(o,o) \times \partial M^*(o,o,\cdot,\cdot)(o,o),$$

in which each "factor" involves both the pair (x,y) of "solution variables" and the pair (z,w) of "Kuhn–Tucker variables."

EXAMPLE 6.14. Take $p = n$ and $q = m$ and define $S(K)$ by $K(u,x,v,y) = \langle u,y\rangle + \langle x,v\rangle$. It is easily checked that $A_i^{-1}(0) \cap (\text{rec cone}_i K) = \{(o,o)\}$ for $i = 1,2$. Hence Lemma 2.9 implies that range $A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$, or in other words $S(L)$ is strongly consistent (see the proof of Lemma 6.4). If $[P]$ denotes the equivalence class containing P_1 and P_2, then $[P] = [AK]$ by Theorem 6.5. Since

$$P_1(u,v) = \begin{cases} o & \text{if } u = o \text{ and } v = o \\ +\infty & \text{if } u = o \text{ and } v \neq o \\ -\infty & \text{if } u \neq o \end{cases}$$

this implies that $\text{dom } AK = \text{dom } P = \{(o,o)\}$, whereas $A \text{ dom } K = R^p \times R^q$. Clearly $S(K)$ is strongly consistent, the set of optimal solutions of $S(K)$ is $R^m \times R^n$, and the set of Kuhn–Tucker vectors for $S(K)$ is $R^p \times R^q$. The Lagrangian of $S(K)$ contains the function

$$M(x,w,y,z) = \sup_u \inf_v \{\langle u,z\rangle + \langle v,w\rangle + K(u,x,v,y)\}$$

$$= \begin{cases} o & \text{if } x + w = o \text{ and } y + z = o \\ +\infty & \text{if } x + w = o \text{ and } y + z \neq o \\ -\infty & \text{if } x + w \neq o \end{cases}$$

Hence the set of saddle-points of the Lagrangian is just $\text{dom } M = \{(x,w) | x + w = o\} \times \{(y,z) | y + z = o\}$. Thus, if (z,w) is any given Kuhn–Tucker vector for $S(K)$, the set

$$\{(x,y) | (x,w,y,z) \text{ is a saddle-point of the Lagrangian}\}$$

equals $\{(-w,-z)\}$ and hence is far from the equalling the set of optimal solutions of $S(K)$. It is of interest to note that the dual program $S(L)$ is given by $L(s,z,t,w) = \langle s,w\rangle + \langle z,t\rangle$ and hence is "identical" with $S(K)$.

In order to describe more fully the duality between the optimal solutions of $S(K)$ and $S(L)$, we introduce another definition. For each x in R^m define the function f_x on R^q by $f_x(v) = \inf \underline{K}(o,x,v,\cdot)$, and for each y in R^n define the function g_y on R^p by $g_y(u) = \sup \overline{K}(u,\cdot,o,y)$. It follows easily from (5.7) that f_x is convex and g_y is concave. An optimal solution (x,y) of $S(K)$ is said to be <u>stable</u> iff the directional derivative function

$$v \to f_x'(o;v) = \lim_{\lambda \downarrow o} \lambda^{-1}(f_x(\lambda v) - f_x(o))$$

is never $-\infty$ and the directional derivative function

$$u \to g_y'(o;u) = \lim_{\lambda \downarrow o} \lambda^{-1}(g_y(\lambda u) - g_y(o))$$

is never $+\infty$. It is not hard to show that (x,y) is an optimal solution of $S(K)$ iff $f_x(o) = g_y(o) \in R$. Hence by (23.1) the directional derivatives mentioned in the definition exist ($+\infty$ and $-\infty$ being allowed as limits). Stable optimal solutions of $S(L)$ are defined similarly, using the functions $h_w(s) = \inf \underline{L}(s,\cdot,o,w)$ and $k_z(t) = \sup \overline{L}(o,z,t,\cdot)$.

LEMMA 6.15. <u>Let (x,y) be an optimal solution of $S(K)$. Then (x,y) is stable iff f_x and g_y are subdifferentiable at the origin.</u>

PROOF. By (23.2) and (23.3).

LEMMA 6.16. <u>For $(x,y) \in R^m \times R^n$ and $(z,w) \in R^p \times R^q$, each of the following conditions is equivalent to the extremality conditions in Lemma 6.11:</u>

(e) $-(z,w) \in \partial g_y(o) \times \partial f_x(o)$ <u>and</u> $f_x(o) = g_y(o) \in R$;

(f) $-(x,y) \in \partial h_w(o) \times \partial k_z(o)$ <u>and</u> $h_w(o) = k_z(o) \in R$.

PROOF. By (37.4) and (37.4.1), $(-z,o,-w,o) \in \partial K(o,x,o,y)$ occurs iff

$$\overline{K}(o,x,o,y) = \underline{K}(o,x,o,y) = \alpha \in R$$

and

$$\langle u,z\rangle + \bar{K}(u,\tilde{x},o,y) \leq \alpha \leq \underline{K}(o,x,v,\tilde{y}) + \langle v,w\rangle$$

for all $u \in R^p$, $\tilde{x} \in R^m$, $v \in R^q$ and $\tilde{y} \in R^n$. But this occurs iff

$$g_y(o) = f_x(o) = \alpha \in R$$

and

$$\langle u,z\rangle + g_y(u) \leq \alpha \leq f_x(v) + \langle v,w\rangle$$

for all $u \in R^p$ and $v \in R^q$. This last condition holds iff (*) holds. Similarly, $(-x,o,-y,o) \in \partial L(o,z,o,w)$ occurs iff (†) holds.

THEOREM 6.17. *The pair* (x,y) *is a stable optimal solution of* $S(K)$ *iff* $(-z,o,-w,o) \in \partial K(o,x,o,y)$ *for some pair* (z,w).

PROOF. By Lemma 6.15, (x,y) is a stable optimal solution of $S(K)$ iff $f_x(o) = g_y(o) \in R$ and $\partial g_y(o) \times \partial f_x(o)$ is nonempty. Now apply Lemma 6.15.

COROLLARY 6.17.1. *The program* $S(K)$ *has a stable optimal solution iff* $S(L)$ *does, in which case the two optimal values are equal.*

PROOF. Apply the theorem to both $S(K)$ and $S(L)$ and use Lemma 6.11. The two optimal values are equal by Corollary 6.10.1.

COROLLARY 6.17.2. *If* $S(K)$ *is strongly consistent, then every optimal solution of* $S(K)$ *is stable.*

PROOF. Suppose $S(K)$ is strongly consistent, and let (x,y) be any optimal solution of $S(K)$. By the dual version of Corollary 6.5.1, there exist points z and w such that $(-x,o,-y,o) \in \partial L(o,z,o,w)$. Hence Lemma 6.11 and the theorem imply (x,y) is stable.

COROLLARY 6.17.3. *A pair* (x,y) *is a stable optimal solution of* $S(K)$ *iff there exists a pair* (z,w) *such that* (x,w,y,z) *is a saddle-point of the Lagrangian, and such a pair* (z,w) *is a Kuhn-Tucker vector for* $S(K)$.

PROOF. By Theorem 6.17, Lemma 6.11 and Theorem 6.12.

COROLLARY 6.17.4. *Assume that the optimal value in* $S(K)$ [illegible] *equals* α. *If either*

$$\lim_{\lambda \downarrow o} \lambda^{-1}(P_1(o,\lambda v) - \alpha) = -\infty$$

<u>for some</u> v <u>or</u>

$$\lim_{\lambda \downarrow o} \lambda^{-1}(P_2(\lambda u,o) - \alpha) = +\infty$$

<u>for some</u> u, <u>then neither</u> $S(K)$ <u>nor</u> $S(L)$ <u>has a stable optimal solution.</u>

PROOF. Suppose $(\bar{x},\bar{y})$ is an optimal solution of $S(K)$. Then $f_{\bar{x}}(o) = g_{\bar{y}}(o) \in R$. Notice that

$$f_{\bar{x}} \leq \sup_x f_x = P_1(o,\cdot) \quad \text{and} \quad P_2(\cdot,o) = \inf_y g_y \leq g_{\bar{y}}.$$

Hence the hypothesis implies either $f'_{\bar{x}}(o;v) = -\infty$ for some v or $g'_{\bar{y}}(o;u) = +\infty$ for some u. This means that $(\bar{x},\bar{y})$ is not stable. Thus $S(K)$ has no stable optimal solution, and by Corollary 6.17.1 neither does $S(L)$.

According to Corollary 6.17.2, if a program is strongly consistent then all of its optimal solutions are stable. The next example shows that in the absence of strong consistency there may exist unstable optimal solutions.

EXAMPLE 6.18. In Example 6.3 take $m = n = 1$ and take $[J]$ to be such that $[J^*]$ contains the closed proper concave-convex function

$$(s,t) \to \begin{cases} -\sqrt{t} & \text{if } s \in [o,1] \text{ and } t \in [o,1] \\ +\infty & \text{if } s \in [o,1] \text{ and } t \notin [o,1] \\ -\infty & \text{if } s \notin [o,1] \end{cases}$$

Clearly $\operatorname{dom} J^* = [o,1] \times [o,1]$, $\underline{J}^*(o,o) = \bar{J}^*(o,o)$, and $\partial J^*(o,o) = \emptyset$. From the analysis in Example 6.3 it follows that $S(L)$ is consistent but fails to be strongly consistent, $S(L)$ has a well-defined primal problem, and (o,o) is the only optimal solution of $S(L)$. If (o,o) were stable, then by Corollary 6.17.1 there would exist a stable optimal solution of $S(K)$. But the set of optimal solutions of $S(K)$ is easily seen to be $\partial J^*(o,o)$, which is empty. Hence (o,o) is unstable.

By Corollary 6.5.1, if the dual program $S(L)$ is strongly consistent then each Kuhn–Tucker vector (z,w) for $S(K)$ "corresponds" to a saddle-point of

the Lagrangian in the sense that $(o,o,o,o) \in \partial M(x,w,y,z)$ for some pair (x,y). Example 6.18 shows that this need not be true when $S(L)$ fails to be strongly consistent.

§7: Ordinary Saddle Programs and Lagrange Multipliers

Suppose that S and T are nonempty convex subsets of R^m and R^n, respectively, and that H is a finite concave-convex function on $S \times T$, $g_1, \ldots, g_p$ are finite concave functions on S, and $f_1, \ldots, f_q$ are finite convex functions on T. Consider the problem of finding the saddle-points of H with respect to the pairs (x,y) in $S \times T$ satisfying the constraints

$$g_i(x) \geq 0, \quad i = 1, \ldots, p$$

and

$$f_j(y) \leq 0, \quad j = 1, \ldots, q.$$

Under suitable regularity assumptions this problem can be cast in the form of a generalized saddle program of a certain type.

Let H be a closed proper concave-convex function on $R^m \times R^n$, for $i = 1, \ldots, p$ let g_i be a closed proper concave function on R^m such that

$$\mathrm{dom}_1 H \subset \mathrm{dom}\, g_i \quad \text{and} \quad \mathrm{ri}(\mathrm{dom}_1 H) \subset \mathrm{ri}(\mathrm{dom}\, g_i),$$

and for $j = 1, \ldots, q$ let f_j be a closed proper convex function on R^n such that

$$\mathrm{dom}_2 H \subset \mathrm{dom}\, f_j \quad \text{and} \quad \mathrm{ri}(\mathrm{dom}_2 H) \subset \mathrm{ri}(\mathrm{dom}\, f_j).$$

Let subsets $C \subset R^p \times R^m$ and $D \subset R^q \times R^n$ be defined by

$$C = \{(u,x) \mid x \in \mathrm{dom}_1 H \text{ and } g_i(x) \geq u_i \text{ for } i = 1, \ldots, p\}$$

and

$$D = \{(v,y) \mid y \in \mathrm{dom}_2 H \text{ and } f_j(y) \leq v_j \text{ for } j = 1, \ldots, q\},$$

and define a function K on $(R^p \times R^m) \times (R^q \times R^n)$ by

$$K(u,x,v,y) = \begin{cases} H(x,y) & \text{if } (u,x) \in C \text{ and } (v,y) \in D \\ +\infty & \text{if } (u,x) \in C \text{ and } (v,y) \notin D \\ -\infty & \text{if } (u,x) \notin C \end{cases}$$

THEOREM 7.1. *The function K is closed proper concave-convex with domain $C \times D$. Moreover,*

$$\mathrm{ri}\, C = \{(u,x) \mid x \in \mathrm{ri}(\mathrm{dom}_1 H) \text{ and } g_i(x) > u_i \text{ for } i = 1, \ldots, p\}$$

<u>and</u>

$$\text{cl } C = \{(u,x) \mid x \in \text{cl}(\text{dom}_1 H) \text{ and } g_i(x) \geq u_i \text{ for } i = 1,\dots,p\},$$

<u>and similar formulas hold for</u> ri D <u>and</u> cl D.

PROOF. Define functions $H_0,\dots,H_{p+q}$ on $(R^p \times R^m) \times (R^q \times R^n)$ as follows:

$$H_0(u,x,v,y) = H(x,y)$$

$$H_i(u,x,v,y) = \begin{cases} 0 & \text{if } (x,u_i) \in \text{epi } g_i \\ -\infty & \text{if } (x,u_i) \notin \text{epi } g_i \end{cases} \qquad i = 1,\dots,p$$

$$H_{p+j}(u,x,v,y) = \begin{cases} 0 & \text{if } (y,v_j) \in \text{epi } f_j \\ +\infty & \text{if } (y,v_j) \notin \text{epi } f_j \end{cases} \qquad j = 1,\dots,q$$

Clearly,

$$\text{dom } H_0 = (R^p \times \text{dom}_1 H) \times (R^q \times \text{dom}_2 H)$$

$$\text{dom } H_i = \{(u,x) \mid (x,u_i) \in \text{epi } g_i\} \times (R^q \times R^n) \qquad i = 1,\dots,p$$

$$\text{dom } H_{p+j} = (R^p \times R^m) \times \{(v,y) \mid (y,v_j) \in \text{epi } f_j\} \qquad j = 1,\dots,q$$

and from (34.3) it follows that each H_k is closed and proper. Since

$$\text{ri}(\text{dom } H\ ) \cap \dots \cap \text{ri}(\text{dom } H_{p+q}) \neq \emptyset,$$

Theorem 4.2 implies that $[H_0] + \dots + [H_{p+q}]$ is well-defined, has domain

$$C \times D = \text{dom } H_0 \cap \dots \cap \text{dom } H_{p+q},$$

and contains the function K. The formulas for ri C and cl C (resp. ri D and cl D) follow from (6.5), (7.3) and the fact that epi g_i (resp. epi f_j) is closed.

According to the theorem, $S(K)$ is a generalized saddle program on $R^m \times R^n$ with perturbations in $R^p \times R^q$. We call $S(K)$ the <u>ordinary saddle program</u> associated with $H, g_1,\dots,g_p, f_1,\dots,f_q$.

It will be convenient to introduce the following notation. For any subset S of $R^p \times R^m$ write $S_u = \{x \mid (u,x) \in S\}$ for each $u \in R^p$. Similarly, for any subset T of $R^q \times R^n$ write $T_v = \{y \mid (v,y) \in T\}$ for each $v \in R^q$.

Since the feasible solutions of any generalized saddle program are those pairs (x,y) such that $(o,x,o,y) \in \operatorname{dom} K$, the set of feasible solutions of the ordinary saddle program $S(K)$ is just $C_0 \times D_0$, i.e.

$$\{(x,y) \in \operatorname{dom} H \mid g_1(x) \geq o,\ldots,g_p(x) \geq o \text{ and } f_1(y) \leq o,\ldots,f_q(y) \leq o\}.$$

Recall from the general theory that $S(K)$ is consistent iff $S(K)$ has a feasible solution, i.e. iff $C_0 \times D_0$ is nonempty.

COROLLARY 7.1.1. <u>The program</u> $S(K)$ <u>is strongly consistent iff there exists a pair</u> (x,y) <u>in</u> $ri(\operatorname{dom} H)$ <u>such that</u> $g_1(x) > 0,\ldots,g_p(x) > o$ <u>and</u> $f_1(y) < o,\ldots,f_q(y) < o$. <u>Moreover, this equivalence still holds if "ri" is deleted</u>.

PROOF. The original equivalence assertion is immediate from the formulas for $ri\ C$ and $ri\ D$ given by Theorem 7.1. Now suppose $(x,y) \in \operatorname{dom} H$ is such that $g_1(x) > o,\ldots,g_p(x) > o$ and $f_1(y) < o,\ldots,f_q(y) < o$. Let (x_1,y_1) be any element of $ri(\operatorname{dom} H)$. Then (6.1) and (7.5) imply that, for sufficiently small positive λ, the pair

$$(x_\lambda,y_\lambda) = (1 - \lambda)(x,y) + \lambda(x_1,y_1)$$

is in $ri(\operatorname{dom} H)$ and satisfies $g_1(x_\lambda) > o,\ldots,g_p(x_\lambda) > o$ and $f_1(y_\lambda) < o,\ldots,f_q(y_\lambda) < o$.

If $S(K)$ is strongly consistent, then all the objective functions of $S(K)$ are equivalent and hence the notions of optimal value and optimal solution can be expressed in terms of the single objective function

$$K(o,x,o,y) = \begin{cases} H(x,y) & \text{if } x \in C_0 \text{ and } y \in D_0 \\ +\infty & \text{if } x \in C_0 \text{ and } y \notin D_0 \\ -\infty & \text{if } x \notin C_0 \end{cases}.$$

In this event, the optimal value in $S(K)$ exists and equals α iff

$$\sup_{C_0} \inf_{D_0} H = \inf_{D_0} \sup_{C_0} H = \alpha \in R,$$

and (x,y) is an optimal solution of $S(K)$ iff (x,y) is a saddle-point of

H with respect to $C_0 \times D_0$. The characterizations in the following corollary hold even when $S(K)$ is not strongly consistent.

COROLLARY 7.1.2. *Write* $\mathrm{cl}\, C = C'$ *and* $\mathrm{cl}\, D = D'$. *The optimal value in* $S(K)$ *exists and equals* α *iff*

$$\sup_{C_0} \inf_{D_0'} \underline{H} = \inf_{D_0} \sup_{C_0'} \overline{H} = \alpha \in R.$$

A pair (z,w) *is a Kuhn-Tucker vector for* $S(K)$ *iff the optimal value in* $S(K)$ *exists and equals* α *and*

$$\langle u,z\rangle + \inf_{D_0} \sup_{C_u'} \overline{H} \le \alpha \le \sup_{C_0} \inf_{D_v'} \underline{H} + \langle v,w\rangle,\ \forall v\ \forall u.$$

A pair (x,y) *is an optimal solution of* $S(K)$ *iff*

$$\inf_{D_0'} \underline{H}(x,\cdot) = \sup_{C_0'} \overline{H}(\cdot,y) \in R$$

An optimal solution (x,y) *of* $S(K)$ *is stable iff*

$$\lim_{\lambda \downarrow 0} \lambda^{-1}\left(\inf_{D_{\lambda v}'} \underline{H}(x,\cdot) - \inf_{D_0'} \underline{H}(x,\cdot)\right) > -\infty,\ \forall v$$

and

$$\lim_{\lambda \downarrow 0} \lambda^{-1}\left(\sup_{C_{\lambda u}'} \overline{H}(\cdot,y) - \sup_{C_0'} \overline{H}(\cdot,y)\right) < +\infty,\ \forall u.$$

PROOF. By Theorems 7.1 and 0.1, the least member $\underline{K}$ of $[K]$ is the convex closure of K. Direct computation thus yields

$$\underline{K}(u,x,v,y) = \begin{cases} \underline{H}(x,y) & \text{if } (u,x) \in C \text{ and } (v,y) \in D' \\ +\infty & \text{if } (u,x) \in C \text{ and } (v,y) \notin D' \\ -\infty & \text{if } (u,x) \notin C \end{cases}$$

and hence

$$P_1(u,v) = \sup_{C_u} \inf_{D_v'} \underline{H}.$$

Analogous formulas hold for $\overline{K}$ and P_2. A pair (z,w) is a Kuhn-Tucker vector for $S(K)$ iff $P_1(o,o) = P_2(o,o) = \alpha \in R$ and $\langle u,z\rangle + P_2(u,o) \le \alpha \le P_1(o,v) + \langle v,w\rangle, \forall u \forall v$. Since $P_1(o,o) = P_2(o,o) = \alpha \in R$ occurs iff the

optimal value in $S(K)$ exists and equals α, the first two assertions of the corollary follow immediately from the formulas for P_1 and P_2. Now recall from §6 the functions $f_x(v) = \inf \underline{K}(o,x,v,\cdot)$ and $g_y(u) = \sup \overline{K}(u,\cdot,o,y)$. The formulas for $\underline{K}$ and $\overline{K}$ imply that

$$f_x(v) = \inf_{D_v'} \underline{H}(x,\cdot)$$

whenever $x \in C_0$ and

$$g_y(u) = \sup_{C_u'} \overline{H}(\cdot,y)$$

whenever $y \in D_0$. Since a pair (x,y) is an optimal solution of $S(K)$ iff $f_x(o) = g_y(o) \in R$ (and $(y,x) \in C_0 \times D_0$), the third assertion follows immediately. The last assertion also follows immediately from the formulas for f_x and g_y.

All the general theory of §6 can be applied to the ordinary saddle program $S(K)$. However we shall deal only with the question of whether there exists a good Lagrange multiplier principle for $S(K)$. As a first step, the next results identify the Lagrangian and the extremality conditions associated with $S(K)$.

THEOREM 7.2. <u>The Lagrangian of</u> $S(K)$ <u>contains the function</u>

$$(x,w,y,z) \to \begin{cases} H(x,y) + \sum z_i g_i(x) + \sum w_j f_j(y) & \text{if } (x,w) \in S \text{ and } (y,z) \in T \\ +\infty & \text{if } (x,w) \in S \text{ and } (y,z) \notin T \\ -\infty & \text{if } (x,w) \notin S, \end{cases}$$

<u>where</u> $S = \mathrm{dom}_1 H \times R_+^q$ <u>and</u> $T = \mathrm{dom}_2 H \times R_+^p$.

PROOF. By definition, the Lagrangian contains the function

$$\begin{aligned} M(x,w,y,z) &= \sup_u \inf_v \{\langle u,z\rangle + \langle v,w\rangle + K(u,x,v,y)\} \\ &= \sup_{C_x} \inf_{D_y} \{\langle u,z\rangle + \langle v,w\rangle + H(x,y)\}, \end{aligned}$$

where $C_x = \{u \mid (u,x) \in C\}$ and $D_y = \{v \mid (v,y) \in D\}$. Now C_x equals $\{u \mid g_1(x) \geq u_1, \ldots, g_p(x) \geq u_p\}$ when $x \in \mathrm{dom}_1 H$ and equals the empty set

otherwise, and similarly D_y equals $\{v \mid f_1(y) \leq v_1, \ldots, f_q(y) \leq v_q\}$ when $y \in \mathrm{dom}_2 H$ and equals the empty set otherwise. Therefore the conventions imply that $M(x,w,y,z) = -\infty$ when $x \notin \mathrm{dom}_1 H$ and $M(x,w,y,z) = +\infty$ when $x \in \mathrm{dom}_1 H$ and $y \notin \mathrm{dom}_2 H$. When $(x,y) \in \mathrm{dom}\, H$,

$$M(x,w,y,z) = H(x,y) + \sup_{C_x} \{\langle u,z\rangle + \inf_{D_y} \langle v,w\rangle\}$$

$$= \begin{cases} -\infty & \text{if } w \notin R^q_+ \\ +\infty & \text{if } w \in R^q_+ \text{ and } z \notin R^p_+ \\ H(x,y) + \sum z_i g_i(x) + \sum w_j f_j(y) & \text{if } w \in R^q_+ \text{ and } z \in R^p_+. \end{cases}$$

it is easy to show that $\underline{M} = \mathrm{cl}_2 M$ is given by

$$\underline{M}(x,w,y,z) = \begin{cases} \underline{H}(x,y) + \sum z_i g_i(x) + \sum w_j f_j(y) & \text{if } (x,w) \in S \text{ and } (y,z) \in \mathrm{cl}\, T \\ +\infty & \text{if } (x,w) \in S \text{ and } (y,z) \notin \mathrm{cl}\, T \\ -\infty & \text{if } (x,w) \notin S \end{cases}$$

Finally, observe that the function in the theorem is bounded below by $\underline{M}$ and above by M.

COROLLARY 7.2.1. *Two pairs (x,y) and (z,w) satisfy the extremality conditions associated with $S(K)$ iff*

$$(x,y) \in C_0 \times D_0, \quad (z,w) \in R^p_+ \times R^q_+,$$

$$z_i g_i(x) = 0 \quad \text{for } i = 1,\ldots,p,$$

$$w_j f_j(y) = 0 \quad \text{for } j = 1,\ldots,q,$$

$$0 \in \partial_1 H(x,y) + \sum \partial(z_i g_i)(x),$$

$$0 \in \partial_2 H(x,y) + \sum \partial(w_j f_j)(y).$$

The term $\sum \partial(z_i g_i)(x)$ can be replaced by $\sum z_i \partial g_i(x)$, where the summation extends only over those i such that $z_i > 0$. Similarly, the term $\sum \partial(w_j f_j)(y)$ can be replaced by $\sum w_j \partial f_j(y)$, where the summation extends only over those j such that $w_j > 0$.

PROOF. By definition, (x,y) and (z,w) satisfy the extremality conditions iff (x,w,y,z) is a saddle point of the Lagrangian. By the theorem

optimal value in $S(K)$ exists and [illegible] α, the first two assertions of the corollary follow immediately from the formulas for P_1 and P_2. Now recall from §6 the functions $f_x(v) = \inf \underline{K}(o,x,v,\cdot)$ and $g_y(u) = \sup \overline{K}(u,\cdot,o,y)$. The formulas for $\underline{K}$ and $\overline{K}$ imply that

$$f_x(v) = \inf_{D_v'} \underline{H}(x,\cdot)$$

whenever $x \in C_0$ and

$$g_y(u) = \sup_{C_u'} \overline{H}(\cdot,y)$$

whenever $y \in D_0$. Since a pair (x,y) is an optimal solution of $S(K)$ iff $f_x(o) = g_y(o) \in R$ (and $(y,x) \in C_0 \times D_0$), the third assertion follows immediately. The last assertion also follows immediately from the formulas for f_x and g_y.

All the general theory of §6 can be applied to the ordinary saddle program $S(K)$. However we shall deal only with the question of whether there exists a good Lagrange multiplier principle for $S(K)$. As a first step, the next results identify the Lagrangian and the extremality conditions associated with $S(K)$.

THEOREM 7.2. <u>The Lagrangian of</u> $S(K)$ <u>contains the function</u>

$$(x,w,y,z) \to \begin{cases} H(x,y) + \Sigma z_i g_i(x) + \Sigma w_j f_j(y) & \text{if } (x,w) \in S \text{ and } (y,z) \in T \\ +\infty & \text{if } (x,w) \in S \text{ and } (y,z) \notin T \\ -\infty & \text{if } (x,w) \notin S, \end{cases}$$

<u>where</u> $S = \mathrm{dom}_1 H \times R_+^q$ <u>and</u> $T = \mathrm{dom}_2 H \times R_+^p$.

PROOF. By definition, the Lagrangian contains the function

$$\begin{aligned} M(x,w,y,z) &= \sup_u \inf_v \{\langle u,z\rangle + \langle v,w\rangle + K(u,x,v,y)\} \\ &= \sup_{C_x} \inf_{D_y} \{\langle u,z\rangle + \langle v,w\rangle + H(x,y)\}, \end{aligned}$$

where $C_x = \{u \mid (u,x) \in C\}$ and $D_y = \{v \mid (v,y) \in D\}$. Now C_x equals $\{u \mid g_1(x) \geq u_1,\dots,g_p(x) \geq u_p\}$ when $x \in \mathrm{dom}_1 H$ and equals the empty set

and (36.3), this occurs iff $(x,y) \in \mathrm{dom}\, H$ and $(z,w) \in R_+^p \times R_+^q$,

$$H(x',y) + \sum z_i g_i(x') + \sum w_j' f_j(y) \le H(x,y) + \sum z_i g_i(x) + \sum w_j f_j(y) \quad (1)$$

for all $(x',w') \in \mathrm{dom}_1 H \times R_+^q$, and

$$H(x,y) + \sum z_i g_i(x) + \sum w_j f_j(y) \le H(x,y') + \sum z_i' g_i(x) + \sum w_j f_j(y') \quad (2)$$

for all $(y',z') \in \mathrm{dom}_2 H \times R_+^p$. Taking $z' = z$ in (2) and using (23.8) implies $o \in \partial_2 H(x,y) + \sum \partial(w_j f_j)(y)$. Taking $y' = y$, $z_i' = 1 + z_i$ and $z_k' = z_k$ for $k \neq i$, (2) implies that $o \le g_i(x)$. This holds for each i. But taking $y' = y$ and $z' = o$ in (2) implies $\sum z_i g_i(x) \le o$. Hence $z_i g_i(x) = o$ for each i. Similarly, (1) implies that $f_j(y) \le o$ and $w_j f_j(y) = o$ for each j and $o \in \partial_1 H(x,y) + \sum \partial(z_i g_i)(x)$. This establishes one implication, and the converse is now clear. Now observe that $w_j > o$ trivially implies $\partial(w_j f_j)(y) = w_j \partial f_j(y)$. On the other hand, if $w_j = o$ then $y \in \mathrm{dom}_2 H \subset \mathrm{dom}\, f_j$ implies $\partial(w_j f_j)(y) = \partial\delta(y|\mathrm{dom}\, f_j) \subset \partial\delta(y|\mathrm{dom}_2 H) = o^+ \partial_2 H(x,y)$ and hence $\partial_2 H(x,y) + \partial(w_j f_j)(y) = \partial_2 H(x,y)$. Thus the term $\sum \partial(w_j f_j)(y)$ can be replaced as indicated. The other assertion is proved similarly.

Variables of the sort $z_1,\ldots,z_p$ and $w_1,\ldots,w_q$ appearing in the Lagrangian of $S(K)$ are known traditionally as <u>Lagrange multipliers</u>. Sometimes this term also denotes the particular values of these variables which satisfy certain "extremality conditions" relating to a "Lagrangian function." In this second sense, Lagrange multipliers $(z_1,\ldots,z_p, w_1,\ldots,w_q) = (z,w)$ for an ordinary or generalized saddle program necessarily form a Kuhn-Tucker vector for the program (Theorem 6.12). However, a Kuhn-Tucker vector need not satisfy the extremality conditions, i.e. need not be a Lagrange multiplier. (This behavior can occur if the dual program fails to be strongly consistent. See the remarks following Example 6.18.) In other words, Kuhn-Tucker vectors are defined even when the extremality conditions are not satisfiable. Thus, Kuhn-Tucker vectors (rather than Lagrange multipliers) are the natural "equilibrium

price vectors" for regularized saddle-point problems.

By the general theory of §6, if (x,y) and (z,w) satisfy the extremality conditions then the optimal value in $S(K)$ exists and equals $H(x,y)$, (x,y) is a stable optimal solution of $S(K)$, and (z,w) is a Kuhn-Tucker vector for $S(K)$. In fact, such pairs (x,y) and (z,w) actually satisfy

$$\langle u,z\rangle + \bar{H}(x',y) \leq H(x,y) \leq \underline{H}(x,y') + \langle v,w\rangle$$

for every $(u,x') \in \text{cl } C$ and every $(v,y') \in \text{cl } D$. (cf. Corollary 7.1.2)

The next theorem is the main existence result.

THEOREM 7.3. *If $S(K)$ is strongly consistent, then the extremality conditions can be satisfied whenever the sets*

$$\{x \in \bigcap_{i=1}^{p} \text{rec cone } g_i \mid \inf\{\text{rec } H(\cdot,y)(x) \mid (o,y) \in \text{ri } D\} \geq o\}$$

and

$$\{y \in \bigcap_{j=1}^{q} \text{rec cone } f_j \mid \sup\{\text{rec } H(x,\cdot)(y) \mid (o,x) \in \text{ri } C\} \leq o\}$$

are closed under scaler multiplication by -1.

PROOF. By Corollary 6.17.2 and Theorem 6.17, if $S(K)$ is strongly consistent and has an optimal solution then the extremality conditions can be satisfied. The remainder of the proof consists of showing that Corollary 6.2.1 applies to yield an optimal solution. Suppose $S(K)$ is strongly consistent. By Theorem 5.2, $S(K)$ has a well-defined primal problem which is given by the closed proper equivalence class $[K_0]$. Moreover, $\text{dom } K_0 = C_0 \times D_0$, $\text{ri}(\text{dom } K_0) = (\text{ri } C)_0 \times (\text{ri } D)_0$, and $[K_0]$ contains the function

$$K(x,y) = \begin{cases} H(x,y) & \text{if } x \in C_0 \text{ and } y \in D_0 \\ +\infty & \text{if } x \in C_0 \text{ and } y \notin D_0 \\ -\infty & \text{if } x \notin C \end{cases}.$$

Let $Y = \{y \mid f_1(y) \leq o,\dots,f_q(y) \leq o\}$. Then $D_0 = Y \cap \text{dom}_2 H$, so that $K_0(x,\cdot) = H(x,\cdot) + \delta(\cdot\mid Y)$ whenever $x \in \text{ri}(\text{dom}_1 H)$. It follows from the definitions and (9.3) that

$$(\text{rec}_2 K_0)(y) = \sup\{\text{rec } H(x,\cdot)(y) + \text{rec } \delta(\cdot|Y)(y) \mid (o,x) \in \text{ri } C\}.$$

Now note that $\text{rec } \delta(\cdot|Y) = \delta(\cdot|o^+ Y)$ by (8.5), and $o^+ Y = \bigcap_{j=1}^{q} \text{rec cone } f_j$ by (8.3.3) and (8.7). These facts together imply that $(\text{rec}_2 K_0)(y) \leq o$ iff $y \in \bigcap_{j=1}^{q} \text{rec cone } f_j$ and $\sup\{\text{rec } H(x,\cdot)(y) \mid (o,x) \in \text{ri } C\} \leq o$. A similar argument shows that $(\text{rec}_1 K_0)(x) \geq o$ iff $x \in \bigcap_{i=1}^{p} \text{rec cone } g_i$ and $\inf\{\text{rec } H(\cdot,y)(x) \mid (o,y) \in \text{ri } D\} \geq o$. These two equivalences show that the hypothesis is just what is needed to apply Corollary 6.2.1.

For each $(z,w) \in R_+^p \times R_+^q$ define a function $H_{z,w}$ on $R^m \times R^n$ by

$$H_{z,w}(x,y) = \begin{cases} H(x,y) + \sum z_i g_i(x) + \sum w_j f_j(y) & \text{if } x \in \text{dom}_1 H \text{ and } y \in \text{dom}_2 H \\ +\infty & \text{if } x \in \text{dom}_1 H \text{ and } y \notin \text{dom}_2 H \\ -\infty & \text{if } x \notin \text{dom}_1 H \end{cases}.$$

By Theorem 4.2 it follows easily from the blanket regularity assumptions that $H_{z,w}$ is closed and proper and has the same domain as H. Observe that if M denotes the Lagrangian given in the statement of Theorem 7.2,

$$H_{z,w}(x,y) = M(x,w,y,z)$$

for every $(z,w) \in R_+^p \times R_+^q$ and every $(x,y) \in R^m \times R^n$. If $(z,w) \notin R_+^p \times R_+^q$ put $S_{z,w} = \emptyset$, and if $(z,w) \in R_+^p \times R_+^q$ let $S_{z,w}$ denote the set of pairs (x,y) which are saddle-points of $H_{z,w}$ and which satisfy the conditions

$$g_i(x) \geq o \text{ and } z_i g_i(x) = o \text{ for } i = 1,\dots,p$$

and

$$f_j(y) \leq o \text{ and } w_j f_j(y) = o \text{ for } j = 1,\dots,q.$$

(These conditions together with the condition $(z,w) \in R_+^p \times R_+^q$ are traditionally called complementary slackness conditions.)

For ordinary convex programs there exists a good Lagrange multiplier principle (Theorem 28.1 in [44]). The analogous result for ordinary saddle programs would be the following: "If (z,w) is a Kuhn-Tucker vector for $S(K)$, then $S_{z,w}$ is precisely the set of optimal solutions of $S(K)$." However, the

situation is in general more complicated than this.

LEMMA 7.4. *Two pairs* (x,y) *and* (z,w) *satisfy the extremality conditions iff* $(x,y) \in S_{z,w}$.

PROOF. For any $(z,w) \in R^p_+ \times R^q_+$, a pair (x,y) is a saddle-point of $H_{z,w}$ iff $(x,y) \in \operatorname{dom} H$,

$$H(x',y) + \sum z_i g_i(x') \leq H(x,y) + \sum z_i g_i(x), \forall x' \in \operatorname{dom}_1 H$$

and

$$H(x,y) + \sum w_j f_j(y) \leq H(x,y') + \sum w_j f_j(y'), \forall y' \in \operatorname{dom}_2 H.$$

Now it is an easy exercise to show (using (7.5) and (6.1)) that for any convex function f and any convex set C containing $\operatorname{ri}(\operatorname{dom} f)$, $x^* \in \partial f(x)$ iff

$$f(x') \geq f(x) + \langle x^*, x' - x\rangle, \forall x' \in C.$$

But $\operatorname{ri}(\operatorname{dom}(H(x,\cdot) + \sum w_j f_j)) = \operatorname{ri}(\operatorname{dom}_2 H)$ and $\operatorname{ri}(\operatorname{dom}(h(\cdot,y) + \sum z_i g_i)) = \operatorname{ri}(\operatorname{dom}_1)$ whenever $(z,w) \in R^p_+ \times R^q_+$ and $(x,y) \in \operatorname{dom} H$. Hence it follows from these facts and (23.8) that for $(z,w) \in R^p_+ \times R^q_+$, (x,y) is a saddle-point of $H_{z,w}$ iff $(x,y) \in \operatorname{dom} H$,

$$0 \in \partial_1 H(x,y) + \sum \partial(z_i g_i)(x)$$

and

$$0 \in \partial_2 H(x,y) + \sum \partial(w_j f_j)(y).$$

The lemma follows trivially from this by Corollary 7.2.1 and the definition of $S_{z,w}$.

THEOREM 7.5. *The set of stable optimal solutions of* $S(K)$ *is precisely*

$$\{S_{z,w} | (z,w) \in R^p \times R^q\},$$

and when $S(K)$ *is strongly consistent this set coincides with the set of all optimal solutions of* $S(K)$. *If* $S_{z,w} \neq \emptyset$, *then* (z,w) *is a Kuhn–Tucker vector for* $S(K)$; *the converse holds when the program dual to* $S(K)$ *is strongly consistent.*

PROOF. The assertions follow immediately from Lemma 7.4 and Corollaries

6.17.3, 6.17.2 and 6.5.1.

This means that in general there does not exist a good Lagrange multiplier principle for ordinary saddle programs. In certain circumstances, though, there is an analogue of (28.1).

COROLLARY 7.5.1. *Assume that the program dual to* $S(K)$ *is strongly consistent. If* $S(K)$ *has a unique Kuhn-Tucker vector* (z,w) *(or equivalently, a unique optimal solution* (z,w) *of its dual problem), then the set of stable optimal solutions of* $S(K)$ *is nonempty and equals* $S_{z,w}$.

The next result characterizes dual strong consistency for ordinary saddle programs.

LEMMA 7.6. *The program dual to* $S(K)$ *is strongly consistent iff the two sets*

$$\text{rec cone}_1 H \cap \text{rec cone } g_1 \cap \ldots \cap \text{rec cone } g_p,$$

$$\text{rec cone}_2 H \cap \text{rec cone } f_1 \cap \ldots \cap \text{rec cone } f_q$$

are closed under scaler multiplication by -1.

PROOF. By Lemma 6.4 it suffices to show that

$$(\text{rec}_1 K)(x) \geq 0 \quad \text{iff} \quad x \in \text{rec cone}_1 H \cap \bigcap_{i=1}^{p} \text{rec cone } g_i$$

and

$$(\text{rec}_2 K)(y) \leq 0 \quad \text{iff} \quad y \in \text{rec cone}_2 H \cap \bigcap_{j=1}^{q} \text{rec cone } f_j.$$

Let $H_0, \ldots, H_{p+q}$ be as in the proof of Theorem 7.1 and let $(u,x) \in \text{ri } C$. By (9.3),

$$\text{rec } K(u,x,\cdot,\cdot) = \sum \text{rec } H_k(u,x,\cdot,\cdot).$$

Observe that trivially

$$\text{rec } H_0(u,x,\cdot,\cdot)(v,y) = \text{rec } H(x,\cdot)(y)$$

and

$$\text{rec } H_i(u,x,\cdot,\cdot)(v,y) = 0 \quad \text{for} \quad i = 1,\ldots,p.$$

With the aid of (8.5) and (9.1) it is easy to show that

$$\operatorname{rec} H_{p+j}(u,x,v)(x,y) = \delta((y,v_j)\,|\,\operatorname{epi}(\operatorname{rec} f_j))$$

for $j = 1,\ldots,q$. These facts together imply that $(\operatorname{rec}_2 K)(v,y)$ equals $(\operatorname{rec}_2 H)(y)$ when $(\operatorname{rec} f_j)(y) \le v_j$ for $j = 1,\ldots,q$ and equals $+\infty$ otherwise. This establishes the second equivalence stated above. The first can be proved similarly.

Theorem 7.5 and Corollary 7.5.1 are actually valid for any generalized saddle program, provided that for each $(z,w) \in R^m \times R^n$ the set $S_{z,w}$ is defined to be the set of pairs (x,y) such that (x,y) and (z,w) satisfy the extremality conditions. The proofs go through exactly the same except that the definition of $S_{z,w}$ plays the role of Lemma 7.4.

Throughout this section K is a closed proper concave-convex function on $R^m \times R^n$, L is a closed proper convex-concave function on $R^p \times R^q$, and $A = A_1 \times A_2$ is a linear transformation from $R^m \times R^n$ to $R^p \times R^q$. Define

$$X = \{x \in \mathrm{dom}_1 K \mid A_1 x \in \mathrm{dom}_1 L\},$$

$$Y = \{y \in \mathrm{dom}_2 K \mid A_2 y \in \mathrm{dom}_2 L\},$$

$$Z = \{z \in \mathrm{dom}_1 L^* \mid A_1^* z \in \mathrm{dom}_1 K^*\},$$

$$W = \{w \in \mathrm{dom}_2 L^* \mid A_2^* w \in \mathrm{dom}_2 K^*\},$$

and consider the following pair of minimax problems:

(I) Find the saddle-points of $K - LA$ with respect to $X \times Y$;

(II) Find the saddle-points of $L^* - K^*A^*$ with respect to $Z \times W$.

When $m = p$, $n = q$, A is the identity transformation and L is given by

$$L(x,y) = \begin{cases} 0 & \text{if } x \in R^m_+ \text{ and } y \in R^n_+ \\ +\infty & \text{if } x \notin R^m_+ \text{ and } y \in R^n_+ \\ -\infty & \text{if } y \notin R^n_+, \end{cases}$$

problems (I) and (II) are those considered by Rockafellar in [43]. The results in Lebedev–Tynjanskii [30], Tynjanskii [58], and some of the results in Tynjanskii [57] are improved in this section.

Define a function Φ on $(R^p \times R^m) \times (R^q \times R^n)$ by

$$\Phi(u,x,v,y) = \begin{cases} K(x,y) - L(u + A_1 x, v + A_2 y) & \text{if } (u,x) \in \Gamma \text{ and } (v,y) \in \Delta \\ +\infty & \text{if } (u,x) \in \Gamma \text{ and } (v,y) \notin \Delta \\ -\infty & \text{if } (u,x) \notin \Gamma \end{cases}$$

where

$$\Gamma = \{(u,x) \mid x \in \mathrm{dom}_1 K,\ u + A_1 x \in \mathrm{dom}_1 L\},$$

$$\Delta = \{(v,y) \mid y \in \mathrm{dom}_2 K,\ v + A_2 y \in \mathrm{dom}_2 L\}.$$

LEMMA 8.1. *The function* Φ *is closed proper concave-convex with domain* $\Gamma \times \Delta$, *and*

$$\mathrm{ri}\,\Gamma = \{(u,x) \mid x \in \mathrm{ri}(\mathrm{dom}_1 K),\ u + A_1x \in \mathrm{ri}(\mathrm{dom}_1 L)\},$$

$$\mathrm{ri}\,\Delta = \{(v,y) \mid y \in \mathrm{ri}(\mathrm{dom}_2 K),\ v + A_2y \in \mathrm{ri}(\mathrm{dom}_2 L)\}.$$

PROOF. Trivially, Γ is convex. Define $\Gamma_x = \{u \mid (u,x) \in \Gamma\}$ for each x. Then Γ_x is empty when $x \notin \mathrm{dom}_1 K$ and equals $\mathrm{dom}_1 L - A_1x$ when $x \in \mathrm{dom}_1 K$. Hence (6.8) implies that $(u,x) \in \mathrm{ri}\,\Gamma$ iff $x \in \mathrm{ri}(\mathrm{dom}_1 K)$ and $u \in \mathrm{ri}(\mathrm{dom}_1 L - A_1x)$. This establishes the formula for $\mathrm{ri}\,\Gamma$, and the one for $\mathrm{ri}\,\Delta$ is similar. From these formulas and the fact that K and L are closed and proper, it is not hard to verify (using (34.3)) that Φ has the properties asserted.

By Lemma 8.1, Φ determines a generalized saddle program $S(\Phi)$ on $R^m \times R^n$ with perturbations in $R^p \times R^q$. The formulas for $\mathrm{ri}\,\Gamma$ and $\mathrm{ri}\,\Delta$ imply that $S(\Phi)$ is strongly consistent iff $A\,\mathrm{ri}(\mathrm{dom}\,K)$ meets $\mathrm{ri}(\mathrm{dom}\,L)$, and in this case Theorems 2.3, 4.2 and 6.2 imply that $[K - LA]$ is well-defined and gives the primal problem associated with $S(\Phi)$. By (36.3) this is the same as (I). A generalized saddle program having the form of $S(\Phi)$ is said to be <u>of Fenchel type</u>.

It can be computed as an exercise that the program $S(\Psi)$ dual to $S(\Phi)$ may be given by

$$\Psi(s,z,t,w) = \begin{cases} L^*(z,w) - K^*(s + A_1^*z,\ t + A_2^*w) & \text{if } (s,z) \in \Pi \text{ and } (t,w) \in \Omega \\ -\infty & \text{if } (s,z) \notin \Pi \text{ and } (t,w) \in \Omega \\ +\infty & \text{if } (t,w) \notin \Omega \end{cases}$$

where

$$\Pi = \{(s,z) \mid z \in \mathrm{dom}_1 L^*,\ s + A_1^*z \in \mathrm{dom}_1 K^*\},$$

$$\Omega = \{(t,w) \mid w \in \mathrm{dom}_2 L^*,\ t + A_2^*w \in \mathrm{dom}_2 K^*\}.$$

Thus, the dual of a program of Fenchel type is another program of Fenchel type. Hence $S(\Psi)$ is strongly consistent iff $A^*\mathrm{ri}(\mathrm{dom}\,L^*)$ meets $\mathrm{ri}(\mathrm{dom}\,K^*)$, and in this case $[L^* - K^*A^*]$ is well-defined and gives the primal problem

associated with $S(\Psi)$ (i.e. the dual problem associated with $S(\Phi)$). This problem is the same as (II).

With these facts in mind, it is clear that all the results of §6 can be translated into assertions about problems (I) and (II). In the remainder of this section we illustrate some of this.

A saddle-point of $X - LA$ with respect to $X \times Y$ is called an <u>optimal solution</u> of (I). It is convenient to say an optimal solution of (I) is <u>stable</u> iff it is a stable optimal solution of $S(\Phi)$. Similar definitions are used for (II).

LEMMA 8.2. <u>If</u> $A\ \mathrm{ri}(\mathrm{dom}\ K) \cap \mathrm{ri}(\mathrm{dom}\ L) \neq \emptyset$, <u>then</u>

$$\sup_X \inf_Y K - LA \leq \sup_W \inf_Z L^* - K^*A^* \leq \inf_Z \sup_W L^* - K^*A^* \leq \inf_Y \sup_X K - LA.$$

PROOF. By the dual version of Corollary 6.5.2.

LEMMA 8.3. <u>In order that</u> $A^*\mathrm{ri}(\mathrm{dom}\ L^*) \cap \mathrm{ri}(\mathrm{dom}\ K^*) \neq \emptyset$, <u>it is necessary and sufficient that</u>

$$(\mathrm{rec}_1 K)(x) \geq (\mathrm{rec}_1 L)(A_1 x) \quad \underline{\text{imply}} \quad (\mathrm{rec}_1 K)(-x) \geq (\mathrm{rec}_1 L)(-A_1 x)$$

<u>and</u>

$$(\mathrm{rec}_2 K)(y) \leq (\mathrm{rec}_2 L)(A_2 y) \quad \underline{\text{imply}} \quad (\mathrm{rec}_2 K)(-y) \leq (\mathrm{rec}_2 L)(-A_2 y).$$

PROOF. The lemma will follow from Lemma 6.4, once it is verified that $(\mathrm{rec}_1 \Phi)(o,x) \geq o$ iff $(\mathrm{rec}_1 K)(x) \geq (\mathrm{rec}_1 L)(A_1 x)$ and $(\mathrm{rec}_2 \Phi)(o,y) \leq o$ iff $(\mathrm{rec}_2 K)(y) \leq (\mathrm{rec}_2 L)(A_2 y)$. Only the second equivalence will be checked, as the first is analogous. For each $(u,x) \in \mathrm{ri}\ \Gamma$, it follows from Lemma 8.1, (9.3) and (9.5) that

$$\mathrm{rec}\ \Phi(u,x,\cdot,\cdot)(v,y) = \mathrm{rec}\ K(x,\cdot)(y) - \mathrm{rec}\ L(u + A_1 x,\cdot)(v + A_2 y).$$

Hence, $(\mathrm{rec}_2 \Phi)(o,y) \leq o$ iff $\mathrm{rec}\ K(x,\cdot)(y) \leq \mathrm{rec}\ L(u + A_1 x,\cdot)(A_2 y)$ holds for each $(u,x) \in \mathrm{ri}\ \Gamma$. But this latter condition occurs iff $\mathrm{rec}\ K(x,\cdot)(y) \leq \mathrm{rec}\ L(u,\cdot)(A_2 y)$ holds for each $x \in \mathrm{ri}(\mathrm{dom}_1 K)$ and $u \in \mathrm{ri}(\mathrm{dom}_1 L)$, which occurs iff $(\mathrm{rec}_2 K)(y) \leq (\mathrm{rec}_2 L)(A_2 y)$.

It can be shown that the least member of the Lagrangian of $S(\Phi)$ is the function

$$(x,w,y,z) \to \begin{cases} \underline{K}(x,y) + \underline{L}^*(z,w) - \langle A_1 x, z\rangle - \langle A_2 y, x\rangle & \text{if } (x,y) \in C \text{ and } (z,w) \in \text{cl } D \\ +\infty & \text{if } (x,y) \in C \text{ and } (z,w) \notin \text{cl } D \\ -\infty & \text{if } (x,y) \notin C \end{cases}$$

where $C \times D = (\text{dom}_1 K \times \text{dom}_2 L^*) \times (\text{dom}_2 K \times \text{dom}_1 L^*)$ is the domain of the Lagrangian. From this it follows easily by (36.3), (36.4) and (37.4) that two pairs $(x,y) \in R^m \times R^n$ and $(z,w) \in R^p \times R^q$ satisfy the extremality conditions associated with $S(\Phi)$ and $S(\Psi)$ iff

$$A(x,y) \in \partial L^*(z,w) \quad \text{and} \quad A^*(z,w) \in \partial K(x,y).$$

The next three results are simply translations of Theorem 6.17, Corollary 6.17.1 and Corollary 6.17.2.

THEOREM 8.4. *A pair* (x,y) *is a stable optimal solution of (I) iff there exists a pair* (z,w) *such that*

$$A(x,y) \in \partial L^*(z,w) \quad \text{and} \quad A^*(z,w) \in \partial K(x,y) \tag{*}$$

Dually, a pair (z,w) *is a stable optimal solution of (II) iff there exists a pair* (x,y) *such that* (*) *holds.*

THEOREM 8.5. *Problem (I) has a stable optimal solution iff problem (II) does, in which case the optimal values in (I) and (II) are equal.*

THEOREM 8.6. *If* $A\,\text{ri}(\text{dom } K) \cap \text{ri}(\text{dom } L) \neq \emptyset$, *then every optimal solution of (I) is stable. Dually, if* $A^*\text{ri}(\text{dom } L^*) \cap \text{ri}(\text{dom } K^*) \neq \emptyset$, *then every optimal solution of (II) is stable.*

To go along with these results, we have two existence results. The first is a corollary to the next theorem.

THEOREM 8.7. *Assume* $A\,\text{ri}(\text{dom } K) \cap \text{ri}(\text{dom } L) \neq \emptyset$. *Then*

$$\text{dom}(K - LA)^* \subseteq \text{dom } K^* - A^*\text{dom } L^*,$$

and

$$\mathrm{ri}(\mathrm{dom}\, K^* - A^*\mathrm{dom}\, L^*) \subset \mathrm{dom}(K - LA)^*$$

iff

$$\mathrm{rec}_j(K - LA) = \mathrm{rec}_j K - (\mathrm{rec}_j L)A_j$$

for $j = 1$ *and* 2.

PROOF. The first inclusion follows from Corollaries 4.6.3 and 2.4.2. The second assertion follows from Theorems 4.8 and 2.6, with the help of (6.3.1).

COROLLARY 8.7.1. *Assume* $A\,\mathrm{ri}(\mathrm{dom}\, K) \cap \mathrm{ri}(\mathrm{dom}\, L) \neq \emptyset$, $A^*\mathrm{ri}(\mathrm{dom}\, L^*) \cap \mathrm{ri}(\mathrm{dom}\, K^*) \neq \emptyset$, and

$$\mathrm{rec}_j(K - LA) = \mathrm{rec}_j K - (\mathrm{rec}_j L)A_j$$

for $j = 1$ *and* 2. *Then there exists an optimal solution of* (I).

PROOF. The theorem and (6.3.1) imply that $\mathrm{ri}(\mathrm{dom}(K - LA)^*)$ equals $\mathrm{ri}(\mathrm{dom}\, K^* - A^*\mathrm{dom}\, L^*)$, which equals $\mathrm{ri}(\mathrm{dom}\, K^*) - A^*\mathrm{ri}(\mathrm{dom}\, L^*)$ by (6.6) and (6.6.2). Hence $(0,0) \in \mathrm{ri}(\mathrm{dom}(K - LA)^*)$ iff $A^*\mathrm{ri}(\mathrm{dom}\, L^*) \cap \mathrm{ri}(\mathrm{dom}\, K^*) \neq \emptyset$. Thus, (37.5.3) implies there exists a saddle-point of $[K - LA]$.

A more general existence result is the following.

THEOREM 8.8. *Assume that* $A\,\mathrm{ri}(\mathrm{dom}\, K) \cap \mathrm{ri}(\mathrm{dom}\, L) \neq \emptyset$ *and that the following two conditions are satisfied:*

(a) *If* $\mathrm{rec}\, K(x,\cdot)(y) \leq \mathrm{rec}\, L(A_1x,\cdot)(A_2y)$ *for every* $x \in \mathrm{ri}\, X$, *then* $\mathrm{rec}\, K(x,\cdot)(-y) \leq \mathrm{rec}\, L(A_1x,\cdot)(-A_2y)$ *for every* $x \in \mathrm{ri}\, X$.

(b) *If* $\mathrm{rec}\, K(\cdot,y)(x) \geq \mathrm{rec}\, L(\cdot,A_2y)(A_1x)$ *for every* $y \in \mathrm{ri}\, Y$, *then* $\mathrm{rec}\, K(\cdot,y)(-x) \geq \mathrm{rec}\, L(\cdot,A_2y)(-A_1x)$ *for every* $y \in \mathrm{ri}\, Y$.

Then there exists an optimal solution of (I).

PROOF. Since $[\phi_0] = [K - LA]$, the theorem will follow immediately from Corollary 6.2.1 once it is checked that $\mathrm{rec}_2(K - LA)(y) \leq 0$ iff $\mathrm{rec}\, K(x,\cdot)(y) \leq \mathrm{rec}\, L(A_1x,\cdot)(A_2y)$ for every $x \in \mathrm{ri}\, X$, and that $\mathrm{rec}_1(K - LA)(x) \geq 0$ iff $\mathrm{rec}\, K(\cdot,y)(x) \geq \mathrm{rec}\, L(\cdot,A_2y)(A_1x)$ for every $y \in \mathrm{ri}\, Y$. We show only the first equivalence, as the second is similar. By

Theorem 4.2 and 2.8, $[K - LA]$ has domain $X \times Y$ and contains the function H given by

$$H(x,y) = \begin{cases} K(x,y) - LA(x,y) & \text{if } x \in X \text{ and } y \in Y \\ +\infty & \text{if } x \in X \text{ and } y \notin Y \\ -\infty & \text{if } x \notin X. \end{cases}$$

From this together with (9.3) and (9.5), it follows that

$$\operatorname{rec} H(x,\cdot)(y) = \operatorname{rec} K(x,\cdot)(y) - \operatorname{rec} L(A_1x,\cdot)(A_2y)$$

for every $x \in \operatorname{ri} X$. Since $\operatorname{rec}_2(K - LA)(y) = \sup\{\operatorname{rec} H(x,\cdot)(y) \mid x \in \operatorname{ri} X\}$, the equivalence follows.

Finally, we remark that conditions (a) and (b) above are satisfied for example when $X \times Y$ is bounded.

APPENDIX: Polyhedral Refinements

Polyhedral saddle-functions have much nicer properties than do arbitrary closed saddle-functions. Consequently, many of the results in the thesis admit refinements when some of the saddle-functions involved are polyhedral. These refinements generally involve weakening the hypotheses in either or both of two ways. The first way may be described loosely as follows: if a theorem can be proved using a hypothesis of the form $(C \times D) \cap \mathrm{ri}(\mathrm{dom}\, K) \neq \emptyset$, where K is a closed saddle-function and C and D are convex sets, then the same conclusions (and sometimes even sharper ones) can be obtained from the weaker hypothesis $(C \times D) \cap \mathrm{dom}\, K \neq \emptyset$ when K is actually polyhedral. Refinements of this type rest ultimately on the fact that the main tools from convex function theory which are used in the proofs (e.g. (16.3), (16.4), (23.8) and (23.9)) admit polyhedral refinements of the same type. This covers most of the polyhedral refinements. However the results which are essentially assertions about the existence of saddle-points admit refinements of a different sort. Generally speaking, such results hypothesize conditions of the form

$$(\mathrm{rec}_1 K)(-x) \geq 0 \quad \text{whenever} \quad (\mathrm{rec}_1 K)(x) \geq 0$$

and

$$(\mathrm{rec}_2 K)(-y) \leq 0 \quad \text{whenever} \quad (\mathrm{rec}_2 K)(y) \leq 0.$$

These are dual to the condition $(0,0) \in \mathrm{ri}(\mathrm{dom}\, K^*)$ and hence imply the existence of a saddle-point of K. However, when K is proper and polyhedral, it can be shown (using (23.10)) that $\mathrm{dom}\, \partial K^* = \mathrm{dom}\, K^*$ and hence that K has a saddle-point iff $(0,0) \in \mathrm{dom}\, K^*$. Thus, for K proper and polyhedral, only the weaker conditions of the form

$$(\mathrm{rec}_1 K)(x) \leq 0 \quad \text{for all} \quad x$$

and

$$(\mathrm{rec}_2 K)(y) \geq 0 \quad \text{for all} \quad y$$

are needed.

In addition to the various refinements in Rockafellar [44] concerning polyhedral convex sets and functions, there are several other results of a technical nature which are useful in carrying out the proofs of the polyhedral refinements. One such result is the polyhedral version of Lemma 9.3, which is proved by appealing to (20.1) in place of (11.3). Another is the fact that the recession functions of any closed proper concave-convex function K can be represented as

$$(\mathrm{rec}_1 K)(x) = \inf\{\mathrm{rec}\, K(\cdot,y)(x) \mid y \in D\}$$

and

$$(\mathrm{rec}_2 K)(y) = \sup\{\mathrm{rec}\, K(x,\cdot)(y) \mid x \in C\}$$

for any sets C and D such that $\mathrm{ri}(\mathrm{dom}\, K) \subset C \times D \subset \mathrm{dom}\, K$. A third fact, which is essential for the refinements in the mixed case of §4, is a generalization of (6.5). It is that for any convex sets C_1 and C_2 in R^n satisfying $C_1 \cap \mathrm{ri}\, C_2 \neq \emptyset$, one has $\mathrm{ri}(C_1 \cap C_2) \subset C_1 \cap \mathrm{ri}\, C_2$ and $C_1 \cap \mathrm{cl}\, C_2 \subset \mathrm{cl}(C_1 \cap C_2)$. This can be proved by first using a separation argument based on (11.3) to show that $\mathrm{ri}\, C_0 \cap \mathrm{ri}\, C_2 \neq \emptyset$, where $C_0 = C_1 \cap \mathrm{aff}\, C_2$, and then applying (6.5).

The results in §2 which use the hypothesis $\mathrm{range}\, A \cap \mathrm{ri}(\mathrm{dom}\, K) \neq \emptyset$ can be proved assuming just that $\mathrm{range}\, A \cap \mathrm{dom}\, K \neq \emptyset$ when K is polyhedral. The hypotheses in Theorems 3.2, 3.4 and 3.5 can be weakened considerably. For example, the conclusions of Theorem 3.2 still hold if K is proper polyhedral and $(u,v) \in A\, \mathrm{dom}\, K$ satisfies the conditions

$$\inf\{\mathrm{rec}\, K(\cdot,y)(x) \mid y \in \mathrm{dom}_2 K,\ A_2 y = v\} \leq 0 \quad \text{whenever } A_1 x = 0$$

and

$$\sup\{\mathrm{rec}\, K(x,\cdot)(y) \mid x \in \mathrm{dom}_1 K,\ A_1 x = u\} \geq 0 \quad \text{whenever } A_2 y = 0.$$

The polyhedral refinements for §2 and 3 yield refinements for §4 when all of the saddle-functions $K_1,\ldots,K_s$ are polyhedral. In this case the

hypothesis (*) can be replaced by

$$\operatorname{dom} K_1 \cap \ldots \cap \operatorname{dom} K_s \neq \emptyset.$$

In the mixed case, when for example the saddle-functions $K_1, \ldots, K_r$ are polyhedral but $K_{r+1}, \ldots, K_s$ are not, we can still prove everything with (*) replaced by

$$\operatorname{dom} K_1 \cap \ldots \cap \operatorname{dom} K_r \cap \operatorname{ri}(\operatorname{dom} K_{r+1}) \cap \ldots \cap \operatorname{ri}(\operatorname{dom} K_s) \neq \emptyset.$$

The proofs, however, do not follow from §§2 and 3 by the device of representing $[K_1 + \ldots + K_s]$ as $[KA]$ (cf. Theorem 4.6). Instead, one must carry out proofs parallel to those in §§2 and 3 but which appeal to (20.1) in place of (16.3) and the polyhedral version of (23.8) in place of (23.9).

Concerning §6, define a generalized saddle program $S(K)$ to be <u>polyhedral</u> iff K is polyhedral. It is easy to see that $S(K)$ is polyhedral iff its dual program $S(L)$ is polyhedral. The polyhedral refinements of the first type described above take the form of replacing the hypothesis "$S(K)$ is strongly consistent" by the hypothesis "$S(K)$ is polyhedral and consistent." The refinement of Corollary 6.2.1 combines both general types of refinement: If $S(K)$ is polyhedral and consistent, then it has an optimal solution iff $(\operatorname{rec}_1 K_0)(x) \leq 0$ for all x and $(\operatorname{rec}_2 K_0)(y) \geq 0$ for all y.

The polyhedral refinements for §6 yield refinements for §7 when all of the functions H, $g_1, \ldots, g_p$, $f_1, \ldots, f_q$ are polyhedral and also for §8 when both K and L are polyhedral. The mixed case of §7 does not appear to go through in general. The troublesome spot is establishing a version of Corollary 6.17.2, i.e. establishing the existence of Lagrange multipliers. The mixed case of §8, though, does allow refinements of all the results. The hypothesis "$S(\Phi)$ is strongly consistent," i.e. $\operatorname{ri}(\operatorname{dom} L) \cap A\,\operatorname{ri}(\operatorname{dom} K) \neq \emptyset$, is weakened by replacing $\operatorname{ri}(\operatorname{dom} K)$ with $\operatorname{dom} K$ in the event K is polyhedral and by replacing $\operatorname{ri}(\operatorname{dom} L)$ with $\operatorname{dom} L$ in the event L is polyhedral. Then

everything goes through by appealing to the polyhedral case of §2 and the mixed case of §4.

REFERENCES

[1] K. J. Arrow and L. Hurwicz, "Reduction of constrained maxima to saddle-point problems," in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman, ed., University of California Press, 1956, vol. V, 1-20.

[2] K. J. Arrow, L. Hurwicz, and H. Uzawa, Studies in Linear and Nonlinear Programming, Stanford University Press, Stanford, Calif., [illegible].

[3] A. Auslender, "Recherche des points de selle d'une fonction," in Symposium on Optimization, A. V. Balakrishnan, M. Contensou, B. F. de Veubeke, P. Krée, J. L. Lions, and N. N. Moiseev, eds., Springer-Verlag, New York, 1970, 37-52.

[4] C. Berge, "Sur une convexité régulière non linéaire et ses applications à la théorie des jeux," Bull. Soc. Math. France 82 (1954), 301-315.

[5] L. D. Berkovitz, "A survey of differential games," in Mathematical Theory of Control, A. V. Balakrishnan and L. W. Neustadt, eds., Academic Press, New York, 1967, 342-372.

[6] J. Bram, "The Lagrange multiplier theorem for max-min with several constraints," SIAM J. Appl. Math. 14 (1966), 655-667.

[7] G. W. Brown and J. von Neumann, "Solutions of games by differential equations," in Annals of Mathematics Studies, No. 24 (1950), 73-79.

[8] A. Charnes, "Constrained games and linear programming," Proc. Natl. Acad. Sci. U. S. 39 (1953), 639-641.

[9] J. M. Danskin, "Fictitious play for continuous games," Naval Res. Logist. Quart. 1 (1954), 313-320.

[10] ________, The Theory of Max-Min, Springer-Verlag, New York, 1967.

[11] ________, "An algorithm for the solution of concave-convex games," in Computing Methods in Optimization Problems -2, L. A. Zadeh, L. W. Neustadt, and A. V. Balakrishnan, eds., Academic Press, New York, 1969, 67-76.

[12] G. B. Dantzig, "A proof of the equivalence of the programming problem and the game problem," in Activity Analysis of Production and Allocation, T. C. Koopmans, ed., Wiley, New York, 1951, 330-335.

[13] G. B. Dantzig, E. Eisenberg, and R. W. Cottle, "Symmetric dual nonlinear programs," Pacifac J. Math. 15 (1965), 809-812.

[14] V. F. Dem'janov, "Successive approximations for finding saddle points," Soviet Math. Dokl. 8 (1967), 1350-1353.

[15] ________, "Algorithms for some minimax problems," J. Comput. System Sci. 2 (1968), 342-380.

[16] ________, "Finding saddle points on polyhedra," Soviet Math. Dokl. 11

(1970), 558-561.

[17] J. E. Falk, "Lagrange multipliers and nonlinear programming," J. Math. Anal. Appl. 19 (1967), 141-159.

[18] K. Fan, "Fixed-point and minimax theorems in locally convex topological linear spaces." Proc. Natl. Acad. Sci. U. S. 38 (1952), 121-126.

[19] _______, "Minimax theorems," Proc. Natl. Acad. Sci. U. S. 39 (1953), 42-47.

[20] W. Fenchel, "On conjugate convex functions," Canad. J. Math. 1 (1949), 73-77.

[21] _______, Convex Cones, Sets and Functions, mimeographed lecture notes, Princeton University, Princeton N. J., 1951.

[22] D. Gale, H. W. Kuhn, and A. W. Tucker, "Linear programming and the theory of games," in Activity Analysis of Production and Allocation, T. C. Koopmans, ed., Wiley, New York, 1951, 317-329.

[23] A. Ghouila-Houri, "Le théorème minimax de SION," in Theory of Games, Techniques and Applications, A. Mensch, ed., American Elsevier, New York, 1966, 123-129.

[24] I. L. Glicksberg, "A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points," Proc. Amer Math. Soc. 3 (1952). 170-174.

[25] A. D. Ioffe and V. M. Tikhomirov, "Duality of convex functions and extremum problems," Russian Math. Surveys 23 (1968), no. 6, 53-124.

[26] S. Kakutani, "A generalization of Brouwer's fixed point theorem," Duke Math. J. 8 (1941), 457-459

[27] H. Kneser, "Sur une théorème fondamentale de la théorie des jeux," C. R. Acad. Sci. Paris 234 (1952), 2418-2420.

[28] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, Calif., 1951, 481-492.

[29] L. S. Lasdon, Optimization Theory for Large Systems, MacMillan, 1970.

[30] V. N. Lebedev and N. T. Tynjanskii, "Duality theory for concave-convex games," Soviet Math. Dokl. 8 (1967), 752-756.

[31] O. L. Mangasarian and J. Ponstein, "Minimax and duality in nonlinear programming," J. Math. Anal. Appl. 11 (1965), 504-518.

[32] J.-J. Moreau, "Sur la fonction polaire d'une fonction semicontinue supérieurement," C. R. Acad. Sci. Paris 258 (1964), 1128-1131.

[33] ________, "Théorèmes 'inf-sup'," C. R. Acad. Sci. Paris 258 (1964), 2720-2722.

[34] J. von Neumann, "Zur Theorie der Gesellschaftsspiele," Math. Ann. 100 (1928), 295-320.

[35] H. Nikaidô, "On von Neumann's minimax theorem," Pacifac J. Math. 4 (1954), 65-72.

[36] B. N. Pshenichnyi, "The dual method in extremum problems. I, II," Cybernetics I (1965), no. 3, 91-99, no. 4, 72-79.

[37] J. Robinson, "An iterative method of solving a game," Ann. of Math. (2) 54 (1951), 296-301.

[38] R. T. Rockafellar, Convex Functions and Dual Extremum Problems, thesis, Harvard University, Cambridge, Mass., 1963.

[39] ________, "Minimax theorems and conjugate saddle-functions," Math. Scand. 14 (1964), 151-173.

[40] ________, "Helly's theorem and minima of convex functions," Duke Math. J. 32 (1965), 381-397.

[41] ________, "An extension of Fenchel's duality theorem for convex functions," Duke Math. J. 33 (1966), 81-89.

[42] ________, "Duality and stability in extremum problems involving convex functions," Pacifac J. Math. 21 (1967), 167-187.

[43] ________, "A general correspondence between dual minimax problems and convex programs," Pacifac J. Math. 25 (1968), 597-611.

[44] ________, Convex Analysis, Princeton University Press, Princeton, N. J., 1970.

[45] ________, "Saddle-points and convex analysis," lecture notes, International Summer School on "Mathematical Models of Action and Reaction," Varenna, Italy, June, 1970.

[46] ________, "Monotone operators associated with saddle-functions and minimax problems," in Nonlinear Functional Analysis, Proceedings of Symposia in Pure Mathematics XVIII, Part I, F. E. Browder, ed., Amer. Math. Soc., Providence, R. I., 1970.

[47] ________, "Conjugate convex functions in optimal control and the calculus of variations," J. Math. Anal. Appl. 32 (1970), 174-222.

[48] ________, "Generalized Hamiltonian equations for convex problems of Lagrange," Pacifac J. Math. 33 (1970), 411-427.

[49] Y. Sakawa, "On linear differential games," in Mathematical Theory of Control, A. V. Balakrishnan and L. W. Neustadt, eds., Academic Press, New York, 1967, 373-385.

[50] M. Shiffman, "On the equality min max = max min, and the theory of games," RAND Report RM-243, 1949.

[51] M. Sion, "On general minimax theorems," Pacifac J. Math. 8 (1958), 171-176.

[52] M. Slater, "Lagrange multipliers revisited: a contribution to non-linear programming," Cowles Commission Discussion Paper, Math. 403 (1950).

[53] J. Stoer, "Duality in nonlinear programming and the minimax theorem," Numer. Math. 5 (1963), 371-379.

[54] ________, "Über einen Dualitätssatz der nichtlinearen Programmierung," Numer. Math. 6 (1964), 55-58.

[55] J. Stoer and C. Witzgall, Convexity and Optimization in Finite Dimensions I, Springer-Verlag, New York, 1970.

[56] R. Tremolieres, "Algorithmes pour la résolution de problèmes posés sous la forme 'min.max. f(x,y).' Cas non linéaire," report no. Inf. 6909, Institut de Recherche d'Informatique et d'Automatique, Rocquencourt, May, 1969.

[57] N. T. Tynjanskii, "Conjugate concave-convex functions in linear topological spaces and their saddle-points," Mat. Sb. 7 (1969), 503-531.

[58] ________, "General concave-convex games," Soviet Math. Dokl. 10 (1969), 91-94.

[59] J. Ville, "Sur la théorie générale des jeux où intervient l'habilité des joueurs," in Traité du Calcul des Probabilities et de ses Applications 4 (1938), no. 2, 105-113.

[60] A. Wald, "Generalization of a theorem by v. Neumann concerning zero sum two person games," Ann. of Math. 46 (1945), 281-286.